

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 03:08:33 ; Search time 498.358 Seconds
(without alignments)
9692.234 Million cell updates/sec

Title: US-10-658-691-1

Perfect score: 1137

Sequence: 1 gtggtaaccaaccatcaat.....gcattcaggagcacgttga 1137

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	419.4	36.9	1161	2 AAV68073	AAV68073 DNA encod
2	419.4	36.9	1200	2 AAV68074	AAV68074 DNA encod
3	49	4.3	2000	7 ADA71938	ADA71938 Rice gene
4	46.2	4.1	3851	7 ACA40918	ACA40918 Prokaryot
5	46.2	4.1	5496	7 ACA38681	ACA38681 Prokaryot
6	46.2	4.1	6564	4 AAF31637	AAF31637 Mycobacte
7	46.2	4.1	110000	4 AAI99682	Continuation (33 o
8	46.2	4.1	110000	4 AAI99683	Continuation (33 o
9	43.6	3.8	536	9 ADB68842	ADB68842 Minorit
10	42.8	3.8	2000	7 ADA71938	ADA71938 Rice gene
11	40.6	3.6	2823	7 ACA38373	ACA38373 Prokaryot
12	40.6	3.6	2826	7 ACA40558	ACA40558 Prokaryot
13	40.6	3.6	110000	4 AAI99682	Continuation (21 o
14	40.6	3.6	110000	4 AAI99683	Continuation (21 o
15	40.2	3.5	1830	7 ACA37565	ACA37565 Prokaryot
16	39.8	3.5	109519	5 AAS08693	AAS08693 Micromono
17	39.2	3.4	1308	7 ACA38683	ACA38683 Prokaryot
18	39.2	3.4	2703	7 ACA40727	ACA40727 Prokaryot
19	39.2	3.4	110000	4 AAI99682	Continuation (32 o
20	39.2	3.4	110000	4 AAI99683	Continuation (32 o
21	39	3.4	536	9 ADB68842	ADB68842 Minorit
22	38.6	3.4	690	5 AAH67856	AAH67856 C glutami
23	38.6	3.4	813	4 AAF67978	AAF67978 Corynebac

C	24	38.6	3.4	349980	5	AAH68533	C glutami
	25	38	3.3	840	7	ACA45847	ACA45847 Prokaryot
	26	37.6	3.3	1581	7	ACA27336	ACA27336 Prokaryot
	27	37.4	3.3	2838	2	AAO41087	AAO41087 Insulinom
	28	37.4	3.3	2838	5	AAO4567	AAO4567 Human ins
	29	37.4	3.3	2838	6	ABL65833	ABL65833 Lung canc
	30	37.4	3.3	2838	6	ABL65856	ABL65856 Lung canc
	31	37.4	3.3	2838	6	ABL65390	ABL65390 Lung canc
	32	37.2	3.3	3176	5	AAH81799	AAH81799 Human dif
	33	37.2	3.3	3226	3	AAZ65082	AAZ65082 Membrane-
	34	37.2	3.3	3226	3	AAZ77665	AAZ77665 Human PRO
	35	37.2	3.3	3226	4	AAZ21475	AAZ21475 Human CDN
	36	37.2	3.3	3226	5	AAZ44228	AAZ44228 Human PRO
	37	37.2	3.3	3226	7	ABX77940	ABX77940 Human PRO
	38	37.2	3.3	3226	7	ABX80352	ABX80352 Novel hum
	39	37.2	3.3	3226	7	ACA69258	ACA69258 Human CDN
	40	37.2	3.3	3226	7	ACD24084	ACD24084 Novel hum
	41	37.2	3.3	3226	7	ABX90329	ABX90329 Human sec
	42	37.2	3.3	3226	7	ABX64175	ABX64175 CDNA enco
	43	37.2	3.3	3226	7	ACA67225	ACA67225 CDNA enco
	44	37.2	3.3	3226	7	ACA64397	ACA64397 Novel hum
	45	37.2	3.3	3226	7	ACA03834	ACA03834 CDNA enco

ALIGNMENTS

RESULT 1
AAV68073
ID AAV68073 standard; DNA; 1161 BP.
XX
AC AAV68073;
XX
DT 12-JAN-1999 (first entry)
DE
DE DNA encoding an oxidase.
XX
KW Oxidase; dioxin removal; ds.
XX
OS Pseudomonas sp.
XX
FH Key Location/Qualifiers
FT CDS 1..1161
FT /tag= a
FT /product= "oxidase"
XX
PN JP10257895-A.
XX
PD 29-SEP-1998.
XX
PF 18-MAR-1997; 97JP-00084401.
XX
PR 18-MAR-1997; 97JP-00084401.
XX (ASAH) ASAH KASEI KOGYO KK.
PA AAF31637 Mycobacte
XX
DR WPI; 1998-575906/49.
DR P-PSDB; AAW80331.
XX
XX Oxidase gene derived from a microbe for removal of dioxin - converts hetero:poly:cyclic aromatic hydrocarbon to a benzoic acid homolog.
PT Claim 2; Page 5-6; 15pp; Japanese.
XX
XX The present sequence encodes an oxidase which is derived from a microbe and converts a heteropolycyclic aromatic hydrocarbon to a benzoic acid homologue. The oxidase is used in a method for the removal of dioxin in which dibenzo-p-dioxin is converted to a diphenyl ether triol
XX
SQ Sequence 1161 BP; 305 A; 279 C; 328 G; 249 T; 0 U; 0 Other;
Query Match 36.9%; Score 419.4; DB 2; Length 1161;
Best Local Similarity 62.6%; Pred. No. 6.9e-112;

Matches	689;	Conservative	0;	Mismatches	406;	Indels	6;	Gaps	2;
QY	31	AGAACCAAGGTTTGGAGGCTTATATCGGTGCGAACTCGGGTTCGGAACCAATTGGTAT	90						
Db	37	AGAGTAAAGGCTGGGCGCCCTACGTGGATCGAAGTGGCTTTCGCAATCAITGGTAC	96						
QY	91	CCGGTTCGGCTCGCAGAGGAAATCGCGAAGTACTCCGGTTCGGTCAAGCTCTCTGGGA	150						
Db	97	CCGGTCAATGTTTTCGAAGAGATCGCAGGGCGAGCGGAAGACTAAAGACTGCTCGGT	156						
QY	151	GAGAGATTCTGCTCAATCGGTGGCGGCGAAGTCTATGCGATCCAGACAGAGTGCCTG	210						
Db	157	GAGAACTTCTGCTCATCGTATCGATGGAGCTGTATTGCTCAAGGACGGTGGCTG	216						
QY	211	CATCGGGTGTAAAGCTTTCGACCGGGTGGAGTCTATTCCAGAAACACATATCTCTGC	270						
Db	217	CATCGGGGCTCGAGTGTGCGTCAAGTTCGAGTCAAGGAGTCAAGTCAAGTCAATGC	276						
QY	271	TGGTATCAGGCTGGACATATCGTGGAGCGATGCGCGCTCGTATCTCAAAAC	330						
Db	277	TGGTATCAGGCTGGACATATCGTGGAGCGATGCGCGCTCGTATCTCAAAAC	336						
QY	331	CCCGGAGTGTGAGATCGCGCGGCGCTTTTGAAGAGCTTCCCGGTTGAAGAGCCAAA	390						
Db	337	CCGACAGCGCAGATCGTTCGACAAAGCTGAAACTTTACCAAGTGCAGGAGCCAAAG	396						
QY	391	GGTCTTATCTGTTTACGTAGCGACGCGGAGCAACCAACCGCTTATCGAAGATGTCGG	450						
Db	397	GGTCTGCTTCTTATTTATCTTGGGATGGGAGCGGCTTCTTCCCTTGGCGCGATAGGCA	456						
QY	451	CCCGGCTTCTTATGATGAAACCGCGCATTCACGGCCAAACATCGGCTCGTGGCTCGAAC	510						
Db	457	CCCAATTTCTTGACATGACATGGAATCTCTCGGAGAGAACCAATCATCAAGTCTAAC	516						
QY	511	TGGCGCTTGGTTCGGAACCGCTTTGATGGGGCAGCTTTCATTTCAGAGATTCG	570						
Db	517	TGGCGCTTGGTTCGGAACCGCTTTGATGGGGCAGCTTTCATTTCAGAGATTCG	576						
QY	571	ATCCTGCTGAAGGGCAACGATATCATCTTCCGCTTGGCTTTTCCGCC---TGGCGATGCC	627						
Db	577	ATTCCTGTCAGGACACGATCTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	636						
QY	628	GACAGCTTACGGT---TCCGAGGTTGCTGGGGCAAGCCCAAGGTTTACGATCTG	684						
Db	637	AAGCAACAAACTCGTGTGGTTCAGATGACGTCGTCGAGCGAGGGTGTTCACGATCTT	696						
QY	685	CTTGGCGAGCATTCGGTTCGGGTTTTCGAAGGATCATCGAAGGCAACCTCAATCCAT	744						
Db	697	ATTGGCGACATCGGGTCCAGTGTGGAGGAACTATCGGGGCGAAGTGTCCGCGNA	756						
QY	745	GGCAACATTGGCAGCAAGCGGTCGCCATCAGCATATCGATCTGGTCCGGGCGTACTC	804						
Db	757	GGTGCCTACCGCGGAAATTTGAGGAAACGATATCTCCATTTGGCTCCCGGCTGTCTC	816						
QY	805	AGGTCGAAACCGTGGCGGATCCGAGCTCACGAGTTCGAATGATGATGCGGCTGAT	864						
Db	817	AGGTCGAAACCGTGGCGGATCCGAGCTCACGAGTTCGAATGATGATGCGGCTGAT	876						
QY	865	GAGACGAGCCACTCTACTTTCAGACGCTGGGCAAAAGTCTGTGACGTCAAGGAGCGCA	924						
Db	877	GAAACACACATTTACTTTCCAAACTCTTGGCAAAACCATGTGCCAATGACGAGGAACG	936						
QY	925	GATCTCTCGAGCGAGATTCACGAAATGAGTGGTAGGCTCGCGTTAACGCTTCAT	984						
Db	937	ARGAATTTACGAAACAGAGTTTCGAAAGCAAGTGGTGAACCGATGGCGCTCGAAGGATTC	996						
QY	985	GATGACGACATCATGCGCGTGAATGATGGAGCGCTTCTACGCTGATGATCGGCTTGG	1044						
Db	997	AACGATGACATCTGGGCTCGGAGCTATGTTGATTTCTAGCCGATGATGAAAGCTGG	1056						
QY	1045	TCCGAGAAATCTCTTTCGAGCGGACCGCGCAATCATCGAGTGGGGGGCTTGGCAGT	1104						
Db	1057	GTCAACGAGATTTTGTTCGAGGTGGACGAGGCTATCGTGGCATGGCGCAAGCTGGCGAGC	1116						

QY 1105 CACCAATCGCGCATTCAG 1125
Db 1117 GAACAAATCAGGGTATTCAG 1137

RESULT 2

AAV68074 standard; DNA; 1200 BP.
XX AAV68074;
XX AC
XX AC
DT 12-JAN-1999 (first entry)
XX DNA encoding an oxidase.
XX Oxidase; dioxin removal; ds.
XX Pseudomonas sp.
XX OS
XX FH
XX Key
XX CDS
FT i..1200
FT /tag= a
FT /product= "oxidase"
FT /note= "no termination codon given"

XX JPI0257895-A.
XX 29-SEP-1998.
XX 18-MAR-1997; 97JP-00084401.
XX 18-MAR-1997; 97JP-00084401.
XX (ASAH) ASAH KASEI KOGYO KK.
XX WPI; 1998-575906/49.
XX P-ESDB; AAW80332.

Oxidase gene derived from a microbe for removal of dioxin - converts hetero:poly:cyclic aromatic hydrocarbon to a benzoic acid homolog.

Claim 2; Page 7-8; 15pp; Japanese.

The present sequence encodes an oxidase which is derived from a microbe and converts a heteropolycyclic aromatic hydrocarbon to a benzoic acid homologue. The oxidase is used in a method for the removal of dioxin in which dibenzo-p-dioxin is converted to a diphenyl ether triol

Sequence 1200 BP; 323 A; 284 C; 338 G; 255 T; 0 U; 0 Other;

Query Match 36.9%; Score 419.4; DB 2; Length 1200;
Best Local Similarity 62.6%; Pred. No. 7e-112;
Matches 689; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

QY	31	AGAACCAAGGTTTGGAGGCTTATATCGGTGCGAACTCGGGTTCGGAACCAATTGGTAT	90
Db	94	AGAGTAAAGGCTGGGCGCCCTACGTGGATCGAAGTGGCTTTCGCAATCATTTGGTAC	153
QY	91	CCCGTTCCGCTCGCAGCGAATAATCGCGAAGGTAATCCCGTCCCGTCAAGCTCTCTGGGA	150
Db	154	CCGGTGTATGTTTTCGAAGAGATCGACGAGCGCGAGCAGCACTAAAACCTGCTCGGT	213
QY	151	GAGAGATTTCTGCTCAATCGGTGGCGGCAAGTCTATGCGATCCAGGACAGGTGCTG	210
Db	214	GAGAACTTCTGCTCAATCGTATCGATGGAGCTGATTGCTCAAGGACCGTGGCTG	273
QY	211	CATCGGCTGTAAAGCTTTTCGACCGGCTCGAGTGTCTTATCCAGAAACACATATCTCTGC	270
Db	274	CATCGGCGCTCCAGTTGTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	333
QY	271	TGGTATCAGCGCTGACATATCGTGGGACGATCGCGCTCGTGCATATCTCTCAAAAC	330

334 CYRWRKMRKTKYKRCYATYCCYKRGWYSRRSMETAGKWMRSNRWC 275
545 GCACGCTCTTCATTCACAGATTTCGATCGCTGGTGAAGGCAACGATATCTTGGCGC 604
274 SYSWYMYKWKYKYSYMSYWARSGTWSRGAARTYKYSTSRKWMACRMYACR 215
605 TTGGCTTTGGCTGGCTGGATCCGACGCTTACGCTTCCGAGTTCTCGGGCAAGC 664
214 RYSRTSYCGSYGSSKWKYMSKSMRMTSSCSCYCYCYGAMCWSGMSMYMNGS 155
665 CCAAGGCTTTACGATCTGC 685
154 CCGYTRGKWRKYSKNCCKY 134

RESULT 4
ID ACA0918 standard; DNA; 3851 BP.
AC ACA0918;
XX
19-JUN-2003 (first entry)
XX
Prokaryotic essential gene #22575.
XX
Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
Mycobacterium tuberculosis.
XX
WO200277183-A2.
XX
03-OCT-2002.
XX
21-MAR-2002; 2002WO-US009107.
XX
21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
(ELIT-) ELITRA PHARM INC.
XX
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
WPI; 2003-029926/02.
DR P-PSDB; ABU37049.
XX
New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
Claim 14; SEQ ID NO 28788; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC proliferating proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIFO at
XX ftp.wipo.int/pub/published_pct_sequences

Sequence 3851 BP; 647 A; 1292 C; 1274 G; 635 T; 0 U; 3 Other;
Query Match 4.1%; Score 46.2; DB 7; Length 3851;
Best Local Similarity 46.9%; Pred. No. 0.011;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
QY 614 CGCTGGCGATCCCGACCACTTACGCTTCCGAGTTGTCGGGCAAGCCCAAGGTG 673
Db 1009 CGCTGGAAGAGACCGCGAGGCTACGGTGCCTGTGTGTCGTCGCTGCGCGG 1068
QY 674 TTACGATCTGCTGGCGAGCATTCGTCGCGGTTTCGAAGCATGATCGAAGGCAAC 733
Db 1069 GTGACCGCTGCTGCCCGCGCTGCGATGCTGCTGCGCGACCGCATCAGCCTGAGC 1128
QY 734 CTGCAATCCATGGCAACATTGGCAGCAAGCGCTGCGCATCAGCATATCGATCTGGCTGC 793
Db 1129 TGGCATCCGGCGCGAGCAAGCAGCTGCGCGCGCGATCAGCCGAACTCTCTGGGACG 1188
QY 794 CGGCGCTACTCAAGTTCGACCGTTCGCGCGATCCGAGCTCAGCGAGTTCGATGCTACG 853
Db 1189 GGATCTGTGTGCTGTCTGCGCGCGAGCAAGCAAGTTCGATGCTGCGAGCGCTCAGC 1248
QY 854 TGGCGCTGATGAGACAGCAGCAGCCTCTACTTCCAGACGCTGGGCAAGTCTGTCAGCTCAA 913
Db 1249 TGGAGTTGGCAGCAGCAGCAGCAGCTGCTGTCGCCAGCTGGTTCGAGACGCTGACGGAA 1308
QY 914 AGGAAGC 920
Db 1309 TGGGTGC 1315

RESULT 5
ACA38681
ID ACA38681 standard; DNA; 5496 BP.
XX
AC ACA38681;
XX
19-JUN-2003 (first entry)
XX
Prokaryotic essential gene #20338.
XX
Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
Mycobacterium bovis.
XX
WO200277183-A2.
XX
03-OCT-2002.
XX
21-MAR-2002; 2002WO-US009107.
XX
21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JB, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
DR WPI: 2003-029926/02.
DR P-PSDB; ABU34811.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 26551; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 5496 BP; 941 A; 1859 C; 1795 G; 901 T; 0 U; 0 Other;
Query Match 4.1%; Score 46.2; DB 7; Length 5496;
Best Local Similarity 46.9%; Pred. No. 0.013;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

614 CGCCTGGCGATCCCGACCGCTTACGGCTTCCGAGGTTCCTCGGGCAAGCCCAAGGTG 673
2648 CGCTGGAAAGACCGCCGAGCTAGGGTGCCCTGTGTCTGATCGGTGACCCCGCGG 2707
674 TTACGATCTCTTTGGCAGCATTCGGTGGCGGTTTTCGAGGCGATGATGAGGCAAC 733
2708 GTGACCCGCTGCTGCCCGCGCTGCACTGCTGCTGCGCGACCGCATCAGCATCTCGAGC 2767
734 CTGCAATCCATGCGCAACATTCGGCAGCGCTGCGCATCAGCATATCGATCTGGCTGC 793
2768 TGGCATCCGCGCGCGAAGCCAGCGTCCGCGCGGCGATCAGCGAACCTCTTGGAGC 2827
794 CGGCGGTACTCAAGTTCGAACCGTGGCCGGATCCGAGCTCACGAGTTCGAATGGTACG 853
2828 GGATCGTTGTGTCTGCTCCGCCCGGAGCGAAGCAAGCAATCGATCCGACGAGCTCAAC 2887
854 TGC CGGTGATGAGACGACCGACCTTACTTCCAGACGCTGGGCAAGATGCTGACGTC 913
2888 TGGAGTTGGCACGACACGACGCTGCTGGTGTGCGCAGCGTGTGTCGAGACCGTACGCGAA 2947

QY 914 AGGAAGC 920
DB 2948 TGGGTGC 2954
RESULT 6
AAF31637
ID AAF31637 standard; DNA; 6564 BP.
XX AAF31637;
XX AAF31637;
DT 09-APR-2001 (first entry)
XX Mycobacterium tuberculosis ppsC gene.
XX Mycobacterium tuberculosis; attenuated microorganism;
KW signature tagged transposon mutant; mutant library;
KW mycobacterial infection; actinomycetales; antibacterial; immunostimulant;
KW vaccine; ppsC; ds.
OS Mycobacterium tuberculosis.
XX WO200102555-A1.
XX 11-JAN-2001.
XX 06-JUL-2000; 2000WO-IB000950.
XX 06-JUL-1999; 9US-0142982P.
PR 08-JUL-1999; 99US-0142833P.
XX (INSP) INST PASTEUR.
XX Gicquel B, Guilhot C, Camacho L;
PI WPI: 2001-091804/10.
DR P-PSDB; AAB6467.
XX Screening a mutant library for mutants unable to grow under specific
PT conditions and for identifying loci involved in pathogenicity, comprises
PT using signature tagged transposon mutagenesis.
XX Example 8; Page 122-124; 159pp; English.
XX The present sequence is given in a specification relating to a method for
CC screening a library of mutants. The method comprises constructing a
CC library with insertions in genes and/or regulatory regions of the
CC organisms of interest, where the insertion contains a tag and/or a
CC transposon associated with a tag. The mutants are identified by
CC hybridisation of the tags to known sequences. The method is useful for
CC treating an individual suffering from a mycobacterial infection.
CC suspected of being infected with a Mycobacterium, or having been exposed
CC to an infectious Mycobacterium. It is also useful for identifying and
CC isolating mutants of actinomycetales and for identifying compounds that
CC have antibiotic activity. The method is used to identify mutants of
CC microorganisms, preferably an actinomycetales, such as *M. tuberculosis*,
CC *M. bovis*, *M. leprae*, *M. avium*, *M. intracellulare* and *M.*
CC *paratuberculosis*, that is unable to grow under specific conditions. It is
CC especially useful for identifying loci involved in pathogenicity. It is
CC useful in constructing vaccines. The method can be used to screen
CC multiple libraries concurrently. It can screen libraries of different
CC organisms or different strains of the same organism
XX Sequence 6564 BP; 1137 A; 2189 C; 2154 G; 1084 T; 0 U; 0 Other;
Query Match 4.1%; Score 46.2; DB 4; Length 6564;
Best Local Similarity 46.9%; Pred. No. 0.014;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

614 CGCCTGGCGATCCCGACCGCTTACGGCTTCCGAGGTTCCTCGGGCAAGCCCAAGGTG 673
3716 CGCTGGAAAGACCGCGAGGCTACGGGTGCCCTGTGTCTGATCGGTGACCCCGCGG 3775

QY 674 TTTAGATCTCTTGGGAGCATTCGGTGGCGGTTTTCGAGGAGCATGATCGAAGCAAC 733
Db 3776 GTGACCCGCTGCTGCCCGCTGAGTGTGCTGCGGACCGCATCACGACCTCGAGC 3835
QY 734 CTGCAATCCATGGCAACATTTGGGAGCAAGCGGTCGCCATCAGCATATCATCTGGCTGC 793
Db 3836 TGGCATCGCGCGCCGACCAAGCCACGCTGCGCGCGCGATCAGCGCAACCTCTTGGGACG 3895
QY 794 CGGGCGTACTCAAGGTGCAACCGTGGCGCGGATCCGAGCTCACGAGTTCGAATGGTACG 853
Db 3896 GATCGTTGTGTTGTCGCGCGCGGAGGACGAGCATCGATCGCGGAGGAGCTCAAC 3955
QY 854 TGGCGGTGATGAGACGACCATCTTACTTCCAGACGCTGGGCAAGTGTGATCGTCAA 913
Db 3956 TGGAGTTGGCAGCAGCAGCGAGCTGTGCTGCGCAGCGTGGTGGAGACCGTGAACGGA 4015
QY 914 AGGAAGC 920
Db 4016 TGGGTGC 4022

RESULT 7
AAI99682_32
Continuation (33 of 45) of AAI99682 from base 3200001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682
WP Fragment Name Begin End
WP AAI99682_00 1 110000
WP AAI99682_01 100001 210000
WP AAI99682_02 200001 310000
WP AAI99682_03 300001 410000
WP AAI99682_04 400001 510000
WP AAI99682_05 500001 610000
WP AAI99682_06 600001 710000
WP AAI99682_07 700001 810000
WP AAI99682_08 800001 910000
WP AAI99682_09 900001 1010000
WP AAI99682_10 1000001 1110000
WP AAI99682_11 1100001 1210000
WP AAI99682_12 1200001 1310000
WP AAI99682_13 1300001 1410000
WP AAI99682_14 1400001 1510000
WP AAI99682_15 1500001 1610000
WP AAI99682_16 1600001 1710000
WP AAI99682_17 1700001 1810000
WP AAI99682_18 1800001 1910000
WP AAI99682_19 1900001 2010000
WP AAI99682_20 2000001 2110000
WP AAI99682_21 2100001 2210000
WP AAI99682_22 2200001 2310000
WP AAI99682_23 2300001 2410000
WP AAI99682_24 2400001 2510000
WP AAI99682_25 2500001 2610000
WP AAI99682_26 2600001 2710000
WP AAI99682_27 2700001 2810000
WP AAI99682_28 2800001 2910000
WP AAI99682_29 2900001 3010000
WP AAI99682_30 3000001 3110000
WP AAI99682_31 3100001 3210000
WP AAI99682_32 3200001 3310000
WP AAI99682_33 3300001 3410000
WP AAI99682_34 3400001 3510000
WP AAI99682_35 3500001 3610000
WP AAI99682_36 3600001 3710000
WP AAI99682_37 3700001 3810000
WP AAI99682_38 3800001 3910000
WP AAI99682_39 3900001 4010000
WP AAI99682_40 4000001 4110000
WP AAI99682_41 4100001 4210000
WP AAI99682_42 4200001 4310000
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529

Query Match 4.1%; Score 46.2; DB 4; Length 110000;

Best Local Similarity 46.9%; Pred. No. 0.043;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
QY 614 CGCTGGCGATCCCGACAGCTTACGCTTCGAGTTGCTCGGCGCAAGCCCAAGGTG 673
Db 59397 CGCTGAAAAGACCGCGAGGCTACGGGTGCCCTGTTGCTGATCGGTGACCCCGCGG 59456
QY 674 TTACGATCTGTGTGGCGAGCATTCGCTGCCGCTTTTGAAGGATGATCGAAGCAAC 733
Db 59457 GTGACCCGCTGTGCCCGCTGCACTGCTGCTGCGGACCGCATCACCGACCTCGAGC 59516
QY 734 CTGCAATCCATCGCAACATTTGGCAGCAAGCGGTCGCCATCAGCATATCGATCTGGCTGC 793
Db 59517 TGGCATCCCGGCGGAGCAAGCAACGCTGCGGCGCGGATCAGCGCAACCTCTGGGAGC 59576
QY 794 CGGGCGTACTCAAGGTGCAACCGTGGCGCGGATCCCGAGCTCACGAGTTCGAATGGTACG 853
Db 59577 GGATCGTTGTGTTGTCGCGCGCGGAGCAAGCAATCGATCGCGGACGAGGTCAAC 59636
QY 854 TGGCGGTGATGAGACGACCATCTTACTTCCAGACGCTGGGCAAGTGTGATCGTCAA 913
Db 59637 TGGAGTTGGCAGCAGCAGCGCTGTGCTGCGCAGCGTGGTGGAGACCGGTGACGGA 59696
QY 914 AGGAAGC 920
Db 59697 TGGGTGC 59703

RESULT 8
AAI99683_32
Continuation (33 of 44) of AAI99683 from base 3200001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683
WP Fragment Name Begin End
WP AAI99683_00 1 110000
WP AAI99683_01 100001 210000
WP AAI99683_02 200001 310000
WP AAI99683_03 300001 410000
WP AAI99683_04 400001 510000
WP AAI99683_05 500001 610000
WP AAI99683_06 600001 710000
WP AAI99683_07 700001 810000
WP AAI99683_08 800001 910000
WP AAI99683_09 900001 1010000
WP AAI99683_10 1000001 1110000
WP AAI99683_11 1100001 1210000
WP AAI99683_12 1200001 1310000
WP AAI99683_13 1300001 1410000
WP AAI99683_14 1400001 1510000
WP AAI99683_15 1500001 1610000
WP AAI99683_16 1600001 1710000
WP AAI99683_17 1700001 1810000
WP AAI99683_18 1800001 1910000
WP AAI99683_19 1900001 2010000
WP AAI99683_20 2000001 2110000
WP AAI99683_21 2100001 2210000
WP AAI99683_22 2200001 2310000
WP AAI99683_23 2300001 2410000
WP AAI99683_24 2400001 2510000
WP AAI99683_25 2500001 2610000
WP AAI99683_26 2600001 2710000
WP AAI99683_27 2700001 2810000
WP AAI99683_28 2800001 2910000
WP AAI99683_29 2900001 3010000
WP AAI99683_30 3000001 3110000
WP AAI99683_31 3100001 3210000
WP AAI99683_32 3200001 3310000
WP AAI99683_33 3300001 3410000
WP AAI99683_34 3400001 3510000
WP AAI99683_35 3500001 3610000
WP AAI99683_36 3600001 3710000
WP AAI99683_37 3700001 3810000
WP AAI99683_38 3800001 3910000
WP AAI99683_39 3900001 4010000

CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2823 BP; 478 A; 949 C; 921 G; 475 T; 0 U; 0 Other;

Query Match 3.6%; Score 40.6; DB 7; Length 2823;
Best Local Similarity 49.8%; Pred. No. 0.42;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 346 ATCGCGCGCGCTTTGAAGACGTTCCCGTTGAAGAGCCAAAGCTTATCTTCGTT 405
DB 667 AGCGCGCGGATACGAGTGTCTGCTTGTGCAAGGCCACGCGGCGACTG 726
QY 406 TACGTAGCGCGGCGAACCAACCGCTTATCGAAGATGTGCGCGCGCTTCCTTGAT 465
DB 727 GTGGCGGTGCGCGCGACTTGTGGCGTGTGCGCGCGCGGAGATCGGC 786
QY 466 GAAACCGCGCATTCACGGCCACATCGCTGCTGCGCTGCAACTGCGCTGGGTGCG 525
DB 787 GCTGACGTGCGCTTGGCACCACACACAGGTTGCGAGTGCAGTGGGTTTGGCGCGCG 846
QY 526 GAAACCGCTTTGATGCGCGGCGACGTC 552
DB 847 CATGCGGGTACCTTGGGTGCGCGCG 873

RESULT 12

ACA40558
ID ACA40558 standard; DNA; 2826 BP.

XX ACA40558;
AC

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #22215.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX

OS Mycobacterium tuberculosis.

XX WO20027183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu RH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU3688.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 28428; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2826 BP; 481 A; 949 C; 920 G; 476 T; 0 U; 0 Other;

Query Match 3.6%; Score 40.6; DB 7; Length 2826;
Best Local Similarity 49.8%; Pred. No. 0.42;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 346 ATCGCGCGCGCTTTGAAGACGTTCCCGTTGAAGAGCCAAAGCTTATCTTCGTT 405
DB 667 AGCGCGCGGATACCGACTGTCTGCTTGTGCAAGGCCACGCGCGCGACTG 726
QY 406 TACGTAGCGCGCGAACCAACCGCTTATCGAAGATGTGCGCGCGCTTCCTTGAT 465
DB 727 GTGGCGGTGCGCGCGACTTGTGGCGTGTGCGCGCGGAGATCGGC 786
QY 466 GAAACCGCGCATTCACGGCCACATCGCTGCTGCGCTGCAACTGCGCTGGGTGCG 525
DB 787 GCTGACGTGCGCTTGGCACCACACACAGGTTGCGAGTGCAGTGGGTTTGGCGCGCG 846
QY 526 GAAACCGCTTTGATGCGCGGCGACGTC 552
DB 847 CATGCGGGTACCTTGGGTGCGCGCG 873

RESULT 13

AAI99682.20

Continuation (21 of 45) of AAI99682 from base 2000001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682

WP	Fragment Name	Begin	End
WP	AAI99682.00	1	110000
WP	AAI99682.01	100001	210000
WP	AAI99682.02	200001	310000
WP	AAI99682.03	300001	410000
WP	AAI99682.04	400001	510000
WP	AAI99682.05	500001	610000
WP	AAI99682.06	600001	710000
WP	AAI99682.07	700001	810000
WP	AAI99682.08	800001	910000
WP	AAI99682.09	900001	1010000
WP	AAI99682.10	1000001	1110000
WP	AAI99682.11	1100001	1210000
WP	AAI99682.12	1200001	1310000
WP	AAI99682.13	1300001	1410000
WP	AAI99682.14	1400001	1510000
WP	AAI99682.15	1500001	1610000
WP	AAI99682.16	1600001	1710000
WP	AAI99682.17	1700001	1810000

RESULT 14			Continuation (21 of 44) of AA199683 from base 2000001 (Mycobacterium tuberculosis strain H37Rv) LOCUS AA199683 Accession AA199683	
WP	WP	WP	Sequence split into 44 fragments	LOCUS AA199683 Accession AA199683
WP	WP	WP	Fragment Name	Begin End
WP	AA199683_00	1		110000
WP	AA199683_01	100001		210000
WP	AA199683_02	200001		310000
WP	AA199683_03	300001		410000
WP	AA199683_04	400001		510000
WP	AA199683_05	500001		610000
WP	AA199683_06	600001		710000
WP	AA199683_07	700001		810000
WP	AA199683_08	800001		910000
WP	AA199683_09	900001		1010000
WP	AA199683_10	1000001		1110000
WP	AA199683_11	1100001		1210000
WP	AA199683_12	1200001		1310000
WP	AA199683_13	1300001		1410000
WP	AA199683_14	1400001		1510000
WP	AA199683_15	1500001		1610000
WP	AA199683_16	1600001		1710000
WP	AA199683_17	1700001		1810000
WP	AA199683_18	1800001		1910000

Search completed: June 19, 2004, 04:29:05
Job time : 504.358 secs

	(ELIT-) ELITRA PHARM INC.			
XX	PA			
XX	PB			
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,			
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;			
XX	DR			
XX	WIPI; 2003-029926/02.			
XX	P-PSTDB; ABU33695.			
XX	DR			
PT	New antisense nucleic acids, useful for identifying proteins or screening			
PT	for homologous nucleic acids required for cellular proliferation to			
PT	isolate candidate molecules for rational drug discovery programs.			
XX				
PS	Claim 14; SEQ ID NO 25435; 1766bp; English.			
CC	The invention relates to an isolated nucleic acid comprising any one of			
CC	the 6213 antisense sequences given in the specification where expression			
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:			
CC	(1) a vector comprising a promoter operably linked to the nucleic acid;			
CC	encoding a polypeptide whose expression is inhibited by the antisense			
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated			
CC	polypeptide or its fragment whose expression is inhibited by the			
CC	antisense nucleic acid; (4) an antibody capable of specifically binding			
CC	the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular			
CC	proliferation or the activity of a gene in an organism required for			
CC	proliferation; (7) identifying a compound that influences the activity of			
CC	the gene product or that has an activity against a biological pathway			
CC	required for proliferation, or that inhibits cellular proliferation; (8)			
CC	identifying a gene required for cellular proliferation or the biological			
CC	pathway in which a proliferation-required gene or its gene product lies			
CC	or a gene on which the test compound that inhibits proliferation of an			
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a			
CC	compound's activity; (11) a culture comprising strains in which the gene			
CC	product is overexpressed or underexpressed; (12) determining the extent			
CC	to which each of the strains is present in a culture or collection of			
CC	strains; or (13) identifying the target of a compound that inhibits the			
CC	proliferation of an organism. The antisense nucleic acids are useful for			
CC	identifying proteins or screening for homologous nucleic acids required			
CC	for cellular proliferation to isolate candidate molecules for rational			
CC	drug discovery programs, or for screening homologous nucleic acids			
CC	required for proliferation in cells other than S. aureus, S. typhimurium,			
CC	K. pneumoniae or P. aeruginosa. The present sequence is one of the target			
CC	prokaryotic essential genes. Note: The sequence data for this patent did			
CC	not form part of the printed specification, but was obtained in			
CC	electronic format directly from WIPO at			
CC	fip.wipo.int/pub/published_pct_sequences			
XX				
SQ	Sequence 1830 BP; 338 A; 591 C; 655 G; 246 T; 0 U; 0 Other; Query Match 3.5%; Score 40.2; DB 7; Length 1830; Best Local Similarity 47.4%; Pred. No.0.47; Matches 120; Conservative 0; Mismatches 133; Indels 0; Gaps 0;			
QY	688 GGCGAGCATTCGGTGC CGGTGTTTCGAAGCATTATCGAAGCAAACTGCCATCCATGCCG 747 			
Dd	1306 GACGAGTGCGCGTFCGGGTTCATCGACCGCGGTGTCGGTGCATCCGAAACAACAGTC 1365 			
QY	748 AACATTGGCAGCAAGCGGTGCGCCATCAGCATATCGATCTGGCTGCCGGGTACTCAAG 807 			
Dd	1366 AACCTGGGTGCGGTTCGGACGCGGTGATCATCGCTTCACTGTCGCGCCGAGGCAG 1425 			
QY	808 GTCGAACGGTGGCGGATCCCAGACTCACG CAGTTTCGAATGGTAGTTCGGTTCGATGAG 867 			
Dd	1426 GCCACCAGGTGGCCAACCGCGAGGGGTGGAGATCCGCTACTACTCGGTGATCTACCG 1485 			
QY	868 ACCAGGCACCTCTACTTCCAGAAGCTGGGCAAAGTCGTGCAGTCCAAAGGAGCGGCAGAC 927 			
Dd	1486 GCATCGACGAGATCGAAGAGGCCCTCGCGGCATGCTCAAGCGATTCACGAGAGAAC 1545 			
QY	928 TCCTTCGAGGAG 940 			
Dd	1546 CAGCTGGCGCGC 1558 			


```
Class: sheared ends.
Location/Qualifiers
1. 771
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBE0698K07"
/clone_lib="ZM.0.7.1.5 KB"
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 3.8%; Score 43.4; DB 29; Length 771;
Best Local Similarity 45.8%; Pred. No. 14;
Matches 149; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

Qy 698 CGTGCCGGTTTCGAAGGCATGATCGAAGGAAACCTGCAATCATCGACATTTGCA 757
Db 416 CGAGCGGGTTCACCGCAGAGCGGAGGTGGCCGCTACGCGAGCGGAGC 357
Qy 758 GCAAGCGCTCGCATTCAGATCGATCTGGCTGCGGGCGTACTCAAGGTGCAACCGT 817
Db 356 GCGACGGCGGTTCAGCTGGACTAGCGGTCTGCTGCGACATCGACGCTTCTTCCAGT 297
Qy 818 GCGCGATCCGAGTCAAGCTGCGATGCTGCTGCGCTGCGCTGCGATGAGACCGACC 877
Db 296 CGCCCAAGTGCATGAGCAGACCTACTACTAGCAGCGCTCATGACCCCTCGCGCG 237
Qy 878 TCTACTTCCAGCGCTGGGCAAGTCTGTACGCTCAAGGAGCGGAGAGCTCTTTCGAGC 937
Db 236 CTTTCTTCGCGCGCGCATCGTTGCGGCGAGCAGCAGCGCGCGCGGAT 177
Qy 938 GAGATTCACGAAATGGTAGGCTCGGCTTAAAGCTTCAATGATGAGCATCA 997
Db 176 GTGGAGACGCGGAGCGATGCGCTCTGAGCTTCAGCTTCCCGCTCTGAGTCTGCT 117
Qy 998 TGGCAGTGAATCGATGGAGCGGTT 1022
Db 116 GAGTGAAGCGGCGAGATTCGCT 92

RESULT 2
BX381961 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1072YF05 3-PRIME, mRNA sequence.
ACCESSION BX381961
VERSION BX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D1072CQ3NP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"

FEATURES
source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 3.7%; Score 42.2; DB 13; Length 1201;
Best Local Similarity 3.4%; Pred. No. 34;
Matches 22; Conservative 238; Mismatches 391; Indels 1; Gaps 1;

Qy 369 GTTCCCGGTTGAAGAGCCAAAGTCTTATCTTCGTTTACGTAGGCGAGCGGCAAC 428
Db 415 KKKKCNKNNKKNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 474
Qy 429 GCGCTTATCAAGAGTGTGCGCGCGCTTCCTTGATGAAGAAACCGGCGCATTCACG 488
Db 475 NNNCNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 534
Qy 489 ACATCGGCTCGTGGCTCGAAC-TGGGGCTTGGGTGCGGAAACGGCTTTGATCGGG 547
Db 535 MNAMNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 594
Qy 548 AGCTTCTTCAACAAGATTGATCTGCTGTAAGGCAACGATATCATCTTCTCCGCTG 607
Db 595 MNNNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 654
Qy 608 GTTTGGCGCTGGCGATCCCGACAGCTTACGCTTACGCTTGGAGGTGCTGCGGG 657
Db 655 NNNNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 714
Qy 668 AAGGTGTTACGATCTGCTTGGCGAGCATTCGCTGCGGCTTTTCGAAGGCGATCA 727
Db 715 KKKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 774
Qy 728 GCAACCTGCAATCCATGCAACATTTGGCAGCAAGCGCTCGCCATCAGCATATCG 787
Db 775 NNNNNNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 834
Qy 788 GGCTGCGGGGCTACTCAAGTCAAGCTGCAAGCGTGGCGGATCCGAGCTTCGAT 847
Db 835 MNCNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 894
Qy 848 GGTACGTGCGGCTGATGAGCAGCACCTCTACTTCCAGAGCTGGGCAAAAGTCGTGA 907
Db 895 KNCNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 954
Qy 908 CGTCAAGAGGAGCGGAGACTCTTTCGAGCGAGAAATTCACGAGAAATGGTAGGCT 967
Db 955 MKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 1014
Qy 968 CGCTTAACGGCTCAATGATGAGCAGCATCATGCGACGTGAATCGATCGAGGCC 1019
Db 1015 CANKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 1066

RESULT 3
BX350514/c 1294 bp DNA linear GSS 17-DEC-2002
LOCUS BX350514 pacsl-60.2692.y2 pacsl-60 Pseudomonas aeruginosa genomic clone
DEFINITION pacsl-60.2692, genomic survey sequence.
ACCESSION BX350514
VERSION BX350514.1 GI:27154095
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
```

JOURNAL
COMMENT

J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source

Location/Qualifiers
1. 1294
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone_lib="pacal-60_2692"
/notes="clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

Query Match 3.7%; Score 42.2; DB 28; Length 1294;
Best Local Similarity 45.8%; Pred. No. 35;
Matches 146; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 148 GGAGAGAAGATTCTGCTCAATCGGTGGCGCAAGTCTATCGGATCCAGACAGGTGC 207
DB 496 GGGACAAAGTACTGCGTATGCGGAGCGGTCTCTCTCAAGTGTATCGCCGCAAGCGC 437

QY 208 CTGCATCGCGGTGAACGCTTTCGACCGGGTGCAGTGTCTATTCGAAGAACCATATCG 267
DB 436 CTGGTTTCGCGCGCAGTCCTGTTGCCCTATCAGAAGCGCTTCGACTGCAATGCCAGCGC 377

QY 268 TGCTGGTATCAGCGCTGCAGATATCGCTGGGAGATGCGCGCGCTCGTCGATATCCTCACA 327
DB 376 CTGGCGGGCCGACATGCTCAGCGCAAGTGTGGCTCTGCTGATCCCGCAAG 317

QY 328 AACCCCGGAGTGTGAGATCGCGCGCGCGCTTTGAAGACGTTCCGGTGAAGAGGCC 387
DB 316 CAGCGCTGCACGGTGCAGTTCGACCAAGCTCGCCCTGTAGAACAATGCACGTCGTCAGCCAC 257

QY 388 AAAGGTCTTATCTTCTGTTTACGTAGGCGAGCGGCAACCAACGCGCTTATCGAAGATG 447
DB 256 ACGACCGGAGTACGTCAACGAGTCCATGAGGTCCAGCGGAGTGTATCGCGGATC 197

QY 448 CCGCGCGGCTTCCTTGATG 466
DB 196 CAGCAGCAGCGCTGTACG 178

RESULT 4
BH872653
LOCUS

DEFINITION
hp07e02.b2 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
genomic clone hp07e02 5', genomic survey sequence.

ACCESSION
BH872653
VERSIONKEYWORDS
GSS.

SOURCE

ORGANISM

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 493)

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,

Zutavern, T., McCombie, W.R., and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)

Unpublished (2002)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org

Plate: hp07 row: e column: 02
Seq primer: -21M13UnivFwd
Class: shotgun

High quality sequence stop: 493.

Location/Qualifiers

1. 493

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="hp07e02"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"

/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;

The vector was digested with XbaI and one nucleotide was

added by fill in in the recessive 3' end. The genomic DNA

was nebulized, end repaired, adaptor ligated and size

fractionated using sephadex. The resulting fragments were

between 0.8 and 3 kb and were cloned into the vector

(X/y reads in M13mp19, .b/g reads in pUC19). The same

ligation was transformed in either JM107 or DH5a."

ORIGIN

Query Match 3.6%; Score 40.4; DB 28; Length 493;
Best Local Similarity 51.1%; Pred. No. 60;
Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 372 CCCGTTTGAAGAGGCCAAAGTCTTATCTTCGTTACGTAGGCGACGGCGAACCAAGCC 431
DB 203 CCGGATTCGCGGACCAACGCGCTCCGCGTTCTCGCGCGCGACCAACGAGCGGCC 262

QY 432 GCTTATCGAGATGTCGCGCGCGCTTCTTGATGAAACCGCGCATTCACGGCCAACA 491
DB 263 CGATGTACTCATGTGCGCGCGCGCTCGCGCTTGGTCACGCGCGCGAGGCCGACG 322

QY 492 TCGGCTCGTGGCTCGAACTCGCGCTTGGGTGCGGAAACCGCTTTGATGCGGGGACGT 551
DB 323 TCTCGTGTAGCGCGCGCGACGACGCGTGGTGGCGGAACTCGCGGTTGATCTCGAACTCGG 382

QY 552 CTTTCA 557
DB 383 CGGCGT 388

RESULT 5
BH872654/c
LOCUS

DEFINITION
hp07e02.g2 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
genomic clone hp07e02 5', genomic survey sequence.

ACCESSION
BH872654

VERSION

BH872654.1

KEYWORDS

GSS.

SOURCE

Ze mays

ORGANISM

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 626)

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,

Zutavern, T., McCombie, W.R., and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)

Unpublished (2002)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: hp07 row: e column: 02
 Seq primer: -21M13Univfwd
 Class: shotgun
 High quality sequence stop: 626.
 Location/Qualifiers

FEATURES

1..626
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="hp07e02"
 /lab_host="JM107 or DH5a"
 /clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
 /note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;
 The vector was digested with XbaI and one nucleotide was
 added by fill in the recessive 3' end. The genomic DNA
 was nebulized, end repaired, adaptor ligated and size
 fractionated using sephadex. The resulting fragments were
 between 0.8 and 3 kb and were cloned into the vector
 (x/y reads in M13mp19, b/g reads in pUC19). The same
 ligation was transformed in either JM107 or DH5a."

ORIGIN

Query Match 3.6%; Score 40.4; DB 28; Length 626;
 Best Local Similarity 51.1%; Pred. No. 67;
 Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 372 CCCGTTGAAGAGGCCAAAGTCTTATCTTGTACGTAGCGACGCGCAACCAAGCC 431
 |||||
 Db 533 CCGGATGTCGGGACCAACGCGCTCCGCGTCTCCGCCGTGCGACGACGACGCGCC 474
 |||||
 QY 432 GCTTATCGAAGATGTGCGCGCGCGGCTTCTTGATGAAAACCGCGCATTCACGCGCAACA 491
 |||||
 Db 473 CGATGTACTCCATGTGCGCGCGCGCTCCGCGTGTGTGACGACGCGCGCCGACG 414
 |||||
 QY 492 TCGGCTCGTGGCTCGAAGTGGCGTTCGGTGGCGGAAAACGGCTTTGATCGGGGCAAGT 551
 |||||
 Db 413 TCTCGTGGTGGCGCGCGACGACGCGGTGGTGGCGAACTCCGCGTTGATCTCGAACTCGG 354
 |||||
 QY 552 CTTTCAT 557
 |||||
 Db 353 CGGCGT 348
 |||||

RESULT 6
 LOCUS CC731706 841 bp DNA linear GSS 23-JUN-2003
 DEFINITION OGUD066TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMEM0416L11,
 genomic survey sequence.
 ACCESSION CC731706
 VERSION CC731706.1 GI:32151029
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 841)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGUD066TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers

FEATURES

source

1..841
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMEM0416L11"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 3.6%; Score 40.4; DB 29; Length 841;
 Best Local Similarity 51.1%; Pred. No. 77;
 Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 372 CCCGTTGAAGAGGCCAAAGTCTTATCTTGTACGTAGCGACGCGCAACCAAGCC 431
 |||||
 Db 615 CCGGATGTCGGGACCAACGCGCTCCGCGTCTCCGCCGTGCGACGACGACGCGCC 674
 |||||
 QY 432 GCTTATCGAAGATGTGCGCGCGCGGCTTCTTGATGAAAACCGCGCATTCACGCGCAACA 491
 |||||
 Db 675 CGATGTACTCCATGTGCGCGCGCGCTCCGCGTGTGTGACGACGCGCGCCGACG 734
 |||||
 QY 492 TCGGCTCGTGGCTCGAAGTGGCGTTCGGTGGCGGAAAACGGCTTTGATCGGGGCAAGT 551
 |||||
 Db 735 TCTCGTGGTGGCGCGCGACGACGCGGTGGTGGCGAACTCCGCGTTGATCTCGAACTCGG 794
 |||||
 QY 552 CTTTCAT 557
 |||||
 Db 795 CGGCGT 800
 |||||

RESULT 7

LOCUS CC731699/c 869 bp DNA linear GSS 23-JUN-2003
 DEFINITION OGUD066TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMEM0416L11,
 genomic survey sequence.
 ACCESSION CC731699
 VERSION CC731699.1 GI:32151015
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 869)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGUD066TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers

FEATURES

1..869
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMEM0416L11"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

Query Match

3.6%; Score 40.4; DB 29; Length 869;

```

Best Local Similarity 51.1%; Pred. No. 78;
Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 372 CCGGTTGAAGGCGCAAGGCTTATCTTCGTTACGTAGCGAGCGCGCAACGCC 431
Db 816 CCGGATGTCGGGCGCACCGGCTCCGCTTTCGCCCGTGGCGACACGACGCCAGCC 757
QY 432 GCTTATCGAAGATGTCGCCCGCGCTTCCTTCATGAAACCGCGCAATTCACGGCCAACA 491
Db 756 CGATGTACTCCATGTCGCCCGCGCTCCGCTTGGTGAAGCGCGCCAGGCCGAGCG 697
QY 492 TCGGCTCGGCTCGAATCGCGCTTGGTGGCGGAAACGGCTTGTATGCGGGGACAGT 551
Db 696 TCTCGTGTAGGCGCGGACAGCAGCGTGTGTCGAACTCCGCTTGATCTCGAACTCGG 637
QY 552 CTTTAT 557
Db 636 CCGCGT 631

RESULT 8
BG281668 916 bp mRNA linear EST 21-FEB-2001
LOCUS 602402350F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4544581 5',
DEFINITION mRNA sequence.
ACCESSION BG281668
VERSION BG281668.1 GI:13030594
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC/DICD/BTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW1228 row: C column: 14
High quality sequence stop: 646.
FEATURES
Location/Qualifiers
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4544581"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Size-selected 500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Query Match 3.5%; Score 40; DB 12; Length 916;
Best Local Similarity 44.0%; Pred. No. 18+02;
Matches 166; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 161 TGTCAATCGCGTGGCGCGCAAGGCTATGCGATCCAGACAGGTCCTGTCATCGCGTG 220
Db 228 TGGTAGTCTGCCAGCATCTGTGCCAAGTGGCCCGGACCGCCGTGAAGTAATGCG 287

```

```

QY 221 TAACGCTTCCGACCGGTCGAGTGTATTTCCAGACACCATATCTCTGCTGATCACG 280
Db 288 AGTTCGTGGAGAACTGCAGACTTGGATACTGATTTATGGCTCAGGCTACCCCAATGATC 347
QY 281 GCTGGACATATCGCTGGACGATGGCGCTCGTCGATATCTCACAACCCCGGCGAGT 340
Db 348 CCAAGACAAAGCGTGTGTTGAAGGAGCACGTGGAGCCTGTGTTCGGCTTCCCGCAGTTG 407
QY 341 TGCAGATCGGCGCGCGCTTTGAAGACGTTCCCGTTGAAGAGGCCAAGGCTTATCT 400
Db 408 TCCGGTTCACTGGCGCACGNCAGACCATCTCTGGAGAAAGAGCGGAAGATGTTATAT 467
QY 401 TCGTTTACGTAGCGACGCGGAAACCAACGCGCTTATCGAAGATGTGCCCGCGGCTTC 460
Db 468 GGGAGGACTCAGCATCCGAGATCAGGAGGACTCAGGAAGATCACATCTTCTCTCA 527
QY 461 TTGATGAAACCGCGCATTCAGGCCAATCGGCTCGTGGCTCGAATCGGCTTGG 520
Db 528 ATGAAGGTCCTCCAGCCGCTCCGCTTCTCCAGCGATATTTCTGGAACGCGGCTTGA 587
QY 521 GTGCGGAAACCGGCTTT 537
Db 588 GTCAGCAACCGGCTCT 604

RESULT 9
BX424977/c
LOCUS BX424977 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA005ZH04
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX424977
VERSION BX424977.1 GI:30784421
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 515)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6304.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA005ZH04FPI&cluster=6304.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CLOBA005ZH04FPI.
FEATURES
Location/Qualifiers
1..515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA005ZH04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 3.5%; Score 39.6; DB 13; Length 515;
Best Local Similarity 8.6%; Pred. No. 95;
Matches 16; Conservative 99; Mismatches 71; Indels 0; Gaps 0;

QY 331 CCGGCGAGTGCAGATCGCGCGCGCTTTGAAGACGTTCCCGGTTGAGAGGCCA 390

```


derived from the human host cells."

```

ORIGIN
Query Match      3.4%; Score 39; DB 14; Length 547;
Best Local Similarity 46.2%; Pred. No. 1.4e+02;
Matches 129; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 648 GGTTCCTCGGGCAAGCCCAAGAGGTGTTTACGATCTGCTGGCGAGCATTCGGTCCGGT 707
Db 16 GGTCTCGGGCAAGCATGCTGGCGAGTTCGACCTGGTTGGCATTCCTCGGCTCCGAG 75
QY 708 TTTCAAGGCATGATCGAAGGCAAACTTGCAATCATCGCAACATTTGGCAGCAAGCGCGT 767
Db 76 AGGCGTTCCGAGATCGAAGTTACTTTTCGACGTAGATCGAAGCGGTATCATGAACATCAG 135
QY 768 CGCCATCAGCATATGATCTGCTGCGGGGTACTCTAAGTTCGAACCGTGCAGGATCC 827
Db 136 CGCCGTGCAAGAGACATCTGGCAAGGACAGGAATACGATTCAGTCAAGCGCGGCTT 195
QY 828 CGAGCTCAGCGCATGTTTCAATGTTAGTGCCTGCGGTGATGAGACCACTCTTACTTCCA 887
Db 196 GTCCGACTCACAAATTGAGCAGATGTTCAAGGACGCGAGATGTACAAAGGACGAAGCA 255
QY 888 GACGCTGGCAAGTCGTGACGTCAAGGACGCGCAGA 926
Db 256 GAAGAAGAAGGACGCGGTGACGCGAAGAATGAGGCAGA 294

RESULT 12
LOCUS CA146289 697 bp mRNA linear EST 24-SEP-2003
DEFINITION SCVPR2079804.9 RT2 Saccharum officinarum cDNA clone SCVPR2079804
5', mRNA sequence.
ACCESSION CA146289
VERSION CA146289.1 GI:35045175
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
AUTHORS The libraries that made SUCEST
TITLE Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
JOURNAL Contact: Arruda P
COMMENT Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 079 row: B column: 04
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .697
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCVPR2079804"
/lab_host="DH10B"
/clone_lib="RT2"
/notes="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips (0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(invitrogen). The double-strand cDNAs were fractionated
in a separose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.fad.ic.unicamp.br/public"

ORIGIN
Query Match      3.4%; Score 39; DB 13; Length 697;
Best Local Similarity 45.2%; Pred. No. 1.5e+02;
Matches 141; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 576 GGTGAAGGGCAAGCATATCATTTCTGCCGCTTGCGCTTGGCCCTGGCGATCCCGACAGCT 635
Db 279 GGAGAGAGCGGCGAGATCACATTCGGAGAAATGTCATCTACATCGTCGAGAACATCAACGT 338
QY 636 TAGCGCTTCGAGGTTGCTGCGGGCAAGCCCAAGGTGTTTACGATCTGCTTGGCGAGCA 695
Db 339 CGCGGCCATCGAACCAAGCTCTGTGCTCCATCGAGTGGCGCTCGCAGAGGTTGTCAACCA 398
QY 696 TTGCGTGCCTGGTTTTCGAAGGCGATGATCGAAGCAAACTGCAATCCATCGGCAACATTGG 755
Db 399 CCGCGCGGTGCAGCGCAAGGTCCGCGACGAGATCAAGGCGCTGCTCGGGGACGACGAGCC 458
QY 756 CAGCAAGCGCGTGCCTCAGCATATCGATCTGCTGCGTGGCGGTACTCAAGTCCGAACC 815
Db 459 CATCAAGATGCCACCATCCACAGCTCCCTACCTGCGAGCGGCTCATCAAGAGAGCGT 518
QY 816 GTGCGCCGATCCCGAGCTCACGCGAGTTCGAATGTTAGTTCGCGGTTCGATGAGACGAGCA 875
Db 519 GCGCTCCACTCCCGCATCCGCTCTCTGCTCCGCGACATGAACCTCGAGGAGCCAAAGCT 578
QY 876 CCTTACTTCCA 887
Db 579 CGCGCGGTACNA 590

RESULT 13
LOCUS CB028106 761 bp mRNA linear EST 13-JAN-2003
DEFINITION TgESTzyd47h02.y1 TgRH Tachyzoite Norm 7 cDNA Library toxoplasma
ORGANELAR HEAT SHOCK PROTEIN. ;, mRNA sequence.
ACCESSION CB028106
VERSION CB028106.1 GI:27724478
KEYWORDS EST.
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii
REFERENCE Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,
AUTHORS Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,
Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,
Ritter, E., Bennett, J., Franklin, C., Tsagarishvili, R., Ronko, I.,
Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.
TITLE Toxoplasma EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Clifton, S.
Toxoplasma EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu
Contact David Sibley (toxoeat@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 441.
Location/Qualifiers
1. .761
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH (Type I)"
/db_xref="taxon:5811"
/clone="TgESTzyd47h02.y1"

```



```

FEATURES
    source
        pbloBAC11.
            Location/Qualifiers
                1..645
                    /organism="Drosophila melanogaster"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7227"
                    /clone="BACN08C07"
                    /clone_lib="DrosBAC"
                    /plasmid="pbloBAC11"
                    /notes="end : T7"

```

544	SSANC	SAC	SC	SN	NN	NN	NN	NN	NN	NN	SS	AT	NN	NN	CT	T	A	N	A	N	A	S	S	N	T	G	A	N	G	S	N	S	485																
278	ACG	G	C	T	G	G	A	C	A	T	C	G	T	G	G	A	C	G	A	T	C	T	C	G	A	T	T	C	A	C	A	A	C	C	G	G	C	A	337										
484	N	S	S	A	A	N	S	S	N	A	S	N	T	T	N	S	S	S	C	N	N	N	N	N	S	T	S	N	A	A	A	S	S	S	C	A	S	S	S	A	425								
338	G	T	G	T	C	A	G	A	T	C	G	C	G	C	G	C	T	T	G	A	A	G	A	C	G	T	T	C	C	G	G	T	T	G	A	A	G	G	C	A	397								
424	S	N	A	S	S	S	A	A	S	S	N	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	C	S	C	N	A	S	A	T	S	N	S	N	S	A	365							
398	T	C	T	T	C	G	T	A	G	G	A	C	C	A	C	C	C	G	T	T	A	C	G	A	G	A	T	C	G	A	G	A	T	C	G	G	C	C	C	G	C	456							
364	V	S	N	S	T	N	N	N	N	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	C	A	T	S	N	S	A	S	A	A	A	S	A	S	S	305							
457	T	T	C	T	T	C	A	T	G	A	A	C	C	G	C	A	T	C	A	G	C	A	T	C	G	G	T	C	G	T	C	G	T	C	G	A	A	T	G	G	C	516							
304	N	S	N	S	N	S	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	S	S	S	C	S	C	N	N	S	N	S	S	S	C	C	N	N	N	N	245					
517	T	T	G	G	T	G	C	G	A	A	A	C	G	G	T	T	G	A	T	G	G	G	G	C	A	G	T	T	C	A	T	T	C	A	A	G	A	A	T	T	C	G	A	576					
244	S	S	C	S	S	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	S	S	S	S	S	N	A	A	N	N	N	N	N	N	N	N	N	N	N	185						
577	G	T	G	A	A	G	G	C	A	C	G	A	T	C	A	T	T	C	T	G	C	G	C	T	T	T	T	T	G	C	C	T	T	T	T	G	C	C	T	C	C	G	A	C	C	A	G	T	636
184	C	N	N	N	A	A	C	C	N	A	N	N	N	C	C	C	C	C	C	C	C	C	C	N	C	N	C	A	C	A	C	N	N	N	C	C	C	C	C	C	C	C	C	C	C	C	125		
637	A	C	G	G	T	T	C	C	G	A	G	T	T	C	T	G	C	G	C	G	A	A	C	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	669		
124	A	N	N	N	G	C	C	N	N	N	N	N	N	C	N	C	G	C	M	G	A	C	C	N	N	N	A	92																					

QY 637 ACGGTTCCGAGGTTGTCGGGGCAAGCCAAA 669
| | | | |
Db 124 ANNGGGCCNNNNNNNNNNNCNCGCMGACCNNA 92
| | | | |

RESULT 15	
AJ489090/c	
LOCUS	657 bp mRNA linear EST 35-JUN-2000
DEFINITION	AJ489090 Citrus clementina x Citrus reticulata flavedo Citrus clementina x Citrus reticulata cdNA clone ACR833, mRNA sequence.
ACCESSION	AJ489090
VERSION	AJ489090.1 GI:32241395
KEYWORDS	EST.

3 1 to 645)

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Citrus clementina x Citrus reticulata	Citrus clementina x Citrus reticulata	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Sapindales; Rutaceae; Citrus.	1 (bases 1 to 657)	Gene expression and heat-induced chilling tolerance in Fortune mandarin fruit	Unpublished (2003)	
Citrus clementina	Citrus reticulata		Sanchez-Ballesta, M.		Contact: Sanchez-Ballesta M	
			Sanchez-Ballesta, M.		Ciencia Alimentos-Postcosecha	

tion of this BAC-end sequence was carried out as part of a

AUTHORS	Sanchez-Ballesta, M.
TITLE	Gene expression and heat-induced chilling tolerance in Fortune mandarin fruit
JOURNAL	Unpublished (2003)
COMMENT	Contact: Sanchez-Ballesta M Ciencia Alimentos-Postcosecha

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 04:10:19 ; Search time 114.708 Seconds
(without alignments)
5500.763 Million cell updates/sec

Title: US-10-658-691-1

Perfect score: 1137

Sequence: 1 gtgcctaaccaaccatcaat.....gcattcaggagcagcgttga 1137

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTJUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.2	4.1	4403765	3	US-09-103-840A-2
2	46.2	4.1	4411529	3	US-09-103-840A-1
3	40	3.5	2493	4	US-09-252-991A-12390
4	40	3.5	3024	4	US-09-252-991A-12202
5	39.2	3.4	4403765	3	US-09-103-840A-2
6	39.2	3.4	4411529	3	US-09-103-840A-1
7	38.4	3.4	1053	4	US-09-489-039A-2455
8	37.4	3.3	2838	4	US-08-246-489-1
9	37.2	3.3	6765	4	US-09-620-312D-481
10	37	3.3	891	4	US-09-452-239-1
11	36.4	3.2	975	4	US-09-252-991A-14103
12	36.4	3.2	1869	4	US-09-380-420C-1
13	36.4	3.2	1929	4	US-09-899-642A-1
14	36.4	3.2	1929	4	US-09-899-642A-1
15	36.2	3.2	1314	4	US-09-252-991A-16243
16	36.2	3.2	1500	4	US-09-252-991A-16352
17	36.2	3.2	1584	4	US-09-252-991A-16239
18	36.2	3.2	1758	4	US-09-252-991A-5434
19	36.2	3.2	2013	4	US-09-252-991A-5508
20	35.8	3.1	28958	1	US-08-258-261B-6
21	35.8	3.1	28958	1	US-08-456-837-6
22	35.8	3.1	28958	1	US-08-457-342-6
23	35.8	3.1	28958	1	US-08-457-646A-6
24	35.8	3.1	28958	1	US-08-458-076A-6
25	35.8	3.1	28958	1	US-08-764-232A-4
26	35.8	3.1	28958	1	US-08-457-335A-6
27	35.8	3.1	28958	1	US-08-729-214-6

RESULT 1

US-09-103-840A-2

; Sequence 2, Application US/09103840A

; Patent No. 6294328

; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; TITLE OF INVENTION: TUBERCULOSIS

; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4403765

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; FEATURE:

; OTHER INFORMATION: CDC 1551

; OTHER INFORMATION: "n" bases at various positions throughout the sequence

; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 4.1%; Score 46.2; DB 3; Length 4403765;

Best Local Similarity 46.3%; Pred. No. 0.036; Mismatches 163; Indels 0; Gaps 0;

Matches 144; Conservative 0;

Qy	614	CGCTGGCGATCCCGACCGAGCTTACGCGTTCCGAGGTTGTCGGGCAAGCCCAAGGTG	673
Db	3253673	CGCTGGAAGAAGACCGCGAGGCTACGGGTGCGCTGTGTCGATCGGTACCCGCGCGG	3253732
Qy	674	TTTACGATCTGCTGGCGAGCATTCGGTGGCGGTTTTCGAGGATGATCGAAGGCAAC	733
Db	3253733	GTGACCGGCTGCTGCCCGCGCTGCAGTCGTCGCGACCGCATCACCGACCTCGAGC	3253792
Qy	734	CTGCAATCCATGCAACATTTGGCAGCAAGCGCGTCCGCAATCAGCATATCGATCTGGCTGC	793
Db	3253793	TGCGATCCGCGCGGAGGAGCCACCTGGCGCGGGATCAGCCGACCTCTCTGGGACG	3253852
Qy	794	CGGGCGTACTCAAGGTCGAAACCGTGGCGGATCCCGGATCCGAGTCAGCGAGTTCGATGTCAG	853
Db	3253853	GGATCGTGTGTGCTGTCCGCCCGGAGCAACGACGAATCGATGCCGAGAGGCTCAAC	3253912
Qy	854	TGCGGTCGATGAGACGACGACCTCTACTTCCAGACGCTGGSCAAAGTGTGTCAGCTCAA	913
Db	3253913	TGAGTTGGACCCACGACGACGCTGCTGCTGCCGCGGTGTCGAGACCGTGACCGGA	3253972

ALIGNMENTS

Sequence 6, Appli
Sequence 1, Appli
Sequence 6174, Ap
Sequence 6092, Ap
Sequence 10046, A
Sequence 1369, Ap
Sequence 4086, Ap
Sequence 4070, Ap
Sequence 4119, Ap
Sequence 5746, Ap
Sequence 3821, Ap
Sequence 3777, Ap
Sequence 3690, Ap
Sequence 12, Appli
Sequence 2460, Ap
Sequence 2215, Ap
Sequence 1, Appli
Sequence 2294, Ap

US-09-028-934-6
US-08-764-233A-1
US-09-252-991A-6174
US-09-252-991A-6092
US-09-252-991A-10046
US-09-252-991A-1369
US-09-252-991A-4086
US-09-252-991A-4070
US-09-252-991A-4119
US-09-489-039A-5746
US-09-252-991A-3821
US-09-252-991A-3777
US-09-252-991A-3690
US-09-235-932A-12
US-09-252-991A-2460
US-09-252-991A-2215
US-08-606-322-1
US-09-252-991A-2294

28958 3.1
49377 3.1
564 3.1
696 3.1
1932 3.1
621 3.1
738 3.1
897 3.1
966 3.1
1086 3.1
1986 3.1
2082 3.1
2181 3.1
269 3.1
513 3.1
1155 3.1
2219 3.1
2418 3.1

c 28
c 29
c 30
c 31
c 32
c 33
c 34
c 35
c 36
c 37
c 38
c 39
c 40
c 41
c 42
c 43
c 44
c 45

Qy 914 AGGAAGC 920
|||
Db 3253973 TGGGTGC 3253979

RESULT 2

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match 4.1%; Score 46.2; DB 3; Length 4411529;
Best Local Similarity 46.9%; Pred. No. 0.036;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
Qy 614 CGCTGGCGATCCGACGAGCTTACGGGTTCCGAGGTGTGCGGGCAAGCCCAAGGTG 673
Db 3259397 CGCTGGAAGACCGCGGAGGCTACGGGTGCCCTGTGCTGATCGGTGACCCGCGCG 3259456
Qy 674 TTACGATCTGCTTGGGAGCATTCGGTGGCGGTTTTCGAAGGATGATCGAAGGCAAC 733
Db 3259457 GTGACCGGCTGCTCCGCGGTGCGAGTCGCTGCGCGGACCGCATCACGACCTCGAGC 3259516
Qy 734 CTGCAATCCATGGCAATATTGGCAGCAAGCGCTGCGCATCAGCATATCATCTGGCTGC 793
Db 3259517 TGGCATCGCGCGCGAGCAACGACGCTGCGCGGGGATACCGGACCTCTGGAGC 3259576
Qy 794 CGGCGCTACTAAGGTGGAACCGGTGGCGGATCCGAGCTCAGCGATTCGAATGGTAGC 853
Db 3259577 GGATCGTTGTGTTGTGTCGCGCGCGAGCAACGAGATCGATCCCGGACGAGGCTCAAC 3259636
Qy 854 TGCCGCTCGATGACACGACGACCTTACTTCCAGACGCTGGGCAAGTGTGACGTCRA 913
Db 3259637 TGGAGTTGGACGACGACGACGCTGCTGTGCGCCAGCTGTGTCGAGACCGTACCGGAA 3259696
Qy 914 AGGAAGC 920
|||
Db 3259697 TGGGTGC 3259703

RESULT 3

US-09-252-991A-12390/c
; Sequence 12390, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 12390
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12390

Query Match 3.5%; Score 40; DB 4; Length 2493;
Best Local Similarity 47.0%; Pred. No. 0.081;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 643 TCCGAGGTGTGTCGCGGCAAGCCCAAGGTGTTTACGATCTGCTTGGCGAGCATTCGGTG 702
Db 2355 TCGGGGTGCGCGGTGAGTCGATCCATCCGCTGGGAGCGCGCTGCGCGCAATACCGC 2296
Qy 703 CCGGTTTTGGAAGGATGATCGAAGGCAAACTGCAATCCATGCAACAATGGCAGCAAG 762
Db 2295 GAGGTGCCCCAGGACGATCGCGATTTGGCTGAAGGTTCACTTGGATGCGCCCTCGCGGCCAG 2236
Qy 763 CGCGTCGCGCATCAGCATATCGATCTGGCTGCGCGCGTACTCAAGTTCGAACCGTGGCGC 822
Db 2235 GGCATGCCCATCGTGCCTACGACTACCTGCGCGGAGCGCGGAGGCGCGCGC 2176
Qy 823 GATCCGAGCTCACGAGTTTCGATGTTGATGTCGCGGTGATGAGACGACGACCTCTAC 882
Db 2175 CTCGCTGCGCGCTCGTAGGCAAGGCTACGTGCGGTTCTGTCAGCACCCCGCGCTGAC 2116
Qy 883 TTCCAGACGCTGGGCAAGTCTGTG 906
Db 2115 TACTTGGGGTGAGCGACGTCGAG 2092

RESULT 4

US-09-252-991A-12202
; Sequence 12202, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12202
; LENGTH: 3024
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12202

Query Match 3.5%; Score 40; DB 4; Length 3024;
Best Local Similarity 47.0%; Pred. No. 0.088;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 643 TCCGAGGTGTGTCGCGGCAAGCCCAAGGTGTTTACGATCTGCTTGGCGAGCATTCGGTG 702
Db 814 TCGCGGTGCGCGGTGAGTCGATCCATCCGCTTGGGAGCGCGCTGCGCGCAATACCGC 873
Qy 703 CCGGTTTTGGAAGGATGATCGAAGGCAAACTGCAATCCATGCAACAATTTGGCAGCAAG 762
Db 874 GAGGTGCCCCAGGACGATCGCGATTTGGCTGAAGGTTCACTTGGATGCGCTCGCGGCCAG 933
Qy 763 CGCGTCGCGCATCAGCATATCGATCTGGCTGCGCGCGTACTCAAGTTCGAACCGTGGCGC 822
Db 934 GGCATGCCCATCGTGCCTACGACTACCTGCGCGGAGCGCGGAGGCGCGCGC 993
Qy 823 GATCCGAGCTCACGAGTTTCGATGTTGATGTCGCGGTGATGAGACGACGACCTCTAC 882
Db 994 CTCGCTGCGCGCTCGTAGGAGTCTGTCGCGGTTCTGTCAGCACCCCGCGCTGAC 1053

QY 883 TTCCAGACGCTGGGCAAAAGTCGTG 906
DB 1054 TACCTGGGGTGAGCGACGTCGAG 1077

RESULT 5

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 3.4%; Score 39.2; DB 3; Length 4403765;
Best Local Similarity 45.7%; Pred. No. 4.2; Indels 0; Gaps 0;
Matches 137; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
QY 641 GTTCCGAGGTGCTGGGCAAAAGCTGTTTACGATCTGCTTGGCGGACGATTCGG 700
DB 3139910 GTACCTCGAGGCGCTGGAAGAGCCCTGATGGTATCCAGTGGAGCAGAGTGGTGC 3139851
QY 701 TGGCGGTTTTCGAAGCATGATCGAAGGCAAACTGCAATCCATGCGCAACATTTGGCAGCA 760
DB 3139850 TGGCGGTGATCGACCGCGGTGTCGGCGGATTAACCGAACCACGTCACACCTGGCGTCGG 3139791
QY 761 AGCGCTCGCCATCAGCATATCGATCTGGCTGCGGGCGTACTCAAGGTTCGAACCGTGGC 820
DB 3139790 CTTCGATGGGTGATCATCGTTCATATGTCGGCGCCGAAAGGCAAGCGACCGAGCTGG 3139731
QY 821 CGGATCCGAGCTACGCGAGTTCGAATGATGTCGCGGTGCGGTGCGATGAGACCGACCTCT 880
DB 3139730 CCAGCCGCAAGCGCTGGAGATCCGCTACTACTCGGTCTATCTACCGAGCGATCGACGAGA 3139671
QY 881 ACTTCAGACGCTGGGCAAAAGTCGTGACGTCAAAGGAAGCGCAGACTCTTTCGAGCGAG 940
DB 3139670 TCGAGCAGGCGCTGCTGGCTGCTCAAGCCGATCTACAGGAAACAGCTGGGTGGG 3139611

RESULT 6

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 3.4%; Score 39.2; DB 3; Length 4411529;
Best Local Similarity 45.7%; Pred. No. 4.2;
Matches 137; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 641 GTTCCGAGGTGCTGGGCAAAAGCTGTTTACGATCTGCTTGGCGGACGATTCGG 700
DB 3145742 GTACCTCGAGGCGCTGGAAGAGCCCTGATGGGTATCCAGTGGAGCAGAGTGGTGC 3145683
QY 701 TGGCGGTTTTCGAAGCATGATCGAAGGCAAACTGCAATCCATGGAACCAATTGGCAGCA 760
DB 3145682 TGGCGGTGATCGACCGCGGTGTCGGCGCATACCGAACAACGTCACCTGGCGTCGG 3145623
QY 761 AGCGCTCGCCATCAGCATATCGATCTGGCTGCGGGCGTACTCAAGGTTCGAACCGTGGC 820
DB 3145622 CTTCGATGGGTGATCATCGGTTTCAATGTCGGCGCAAGGCAAGCGACGAGCTGG 3145563
QY 821 CGGATCCGAGCTCACGCGAGTTGGAATGATGACGTGCGGTGCGATGAGACCGACCTCT 880
DB 3145562 CCAGCCGCAAGCGGTGAGATCGCTACTACTCGTCACTACCGAGCGATCGACGAGA 3145503
QY 881 ACTTCAGACGCTGGGCAAAAGTCGTGACGTCAAAGGAAGCGGCGAGACTCTTTCGAGCGAG 940
DB 3145502 TCGAGCAGGCGCTGCTGGCTGCTCAAGCCGATCTACGAGGAAACACGCTGGGTGGG 3145443

RESULT 7

US-09-489-039A-2455
; Sequence 2455, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2455
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2455

Query Match 3.4%; Score 38.4; DB 4; Length 1053;
Best Local Similarity 51.2%; Pred. No. 0.16;
Matches 171; Conservative 0; Mismatches 151; Indels 12; Gaps 3;

QY 85 TGGTATCCCGTTCCGCTCGGAGCGAAATCGCCGAAGTACTCCCGTCCCGTCAAGCTC 144
DB 91 TGGCACCCTGGTGGCGCGCTGCGATATTGCCC---GGCGCCCGCGGTGAAAGCCACCCCTG 147
QY 145 CTGGGAGAGAGATTCTGCTCAATCGCTGGCGGCAAGGTCTATGCGATCCAGGACAGG 204
DB 148 CTGGACGAACAGCTGGTTATCTATCGTATTAAAGCCAGGTGGTGGTGGCCCGGACGTC 207
QY 205 TGCCTGCATCGCGGTGTAAAGCTTTCCGACCGGTCCAGTGCCTATTCCAAAGAACACCATA 264
DB 208 TGCCCGCACCGCGGGTGGCGCTGACCC---TGGGATTTCATGAAGAGGAGGCATC 261
QY 265 TCCTGTGTTGATCAGCGCTGACATATCGCTGGGACGATCGCCGCTCGTGCATCTCTC 324
DB 262 GTCTGCCCTTATCATGCGCTGCGCTTTGGC---GAGGATGGCGCTGCAACCGCATCCCC 318
QY 325 ACAAACCCGCGCAGTGTGCAGATCGCGCGCGCGCTTTGAAGAGCTTCCCGGTTGAAGAG 384

Db 319 TCAGTCCGGGCAACCCATTCGGCCAAACTGCATCTCACCAGCTTCGGCGTGAGGAG 378
Qy 385 GCCAAAGCTTATCTTCGTTTACGTAGGCGAGC 418
Db 379 CGTACGGGCTGATCTGGACCTGCTGCGCTGGC 412

RESULT 8

US-08-246-489-1/c
; Sequence 1, Application US/08246489
; Patent No. 6225049
; GENERAL INFORMATION:
; APPLICANT: Lan, Michael S.
; TITLE OF INVENTION: NOVEL HUMAN INSULINOMA-ASSOCIATED CDNA
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive
; CITY: Newport Beach
; STATE: California
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,489
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/901,715
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH012.012A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2838 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: IA-1

US-08-246-489-1

Query Match 3.3%; Score 37.4; DB 3; Length 2838;
Best Local Similarity 51.5%; Pred. No. 0.5;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 267 CTGCTGTATCAGCGCTGGACATATCGTGGGACGATGGCGGCTCGTCGATATCTCTAC 326
Db 724 CGGCAGGGGCGAGTGGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG 665
Qy 327 AAACCCGGGAGTGTGAGATCGGCGCGGCGGCTTTGAAGACGTTCCCGGTTGAAGAGGC 386
Db 664 ACGCGTGCCTCATCTTGAAGTTCGGCGGCGGCGGAGACGACGGGTGCGCCCGCAGGTGC 605
Qy 387 CAAGAGTCTTATCTTCTGTTTACGAGGCGGCGGACCAACCGCGC 433
Db 604 CGCTTCGCCAGCTCCGCTCG 558

RESULT 9

US-09-620-312D-481
; Sequence 481, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 481
; LENGTH: 6765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (259)..(3105)

US-09-620-312D-481

Query Match 3.3%; Score 37.2; DB 4; Length 6765;
Best Local Similarity 56.6%; Pred. No. 0.85;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 661 AAGCCCAAGGTGTTTACGATCTGTTGGCGAGCATTCGGTGGCGGTTTTCGAAGGCATG 720
Db 892 AGGCACCTAGCATCTCCAATATGCCATGTGTGAATCTGTGACTGTGTCTGAAGACTC 951
Qy 721 ATCGAAGGCAACCTGCAATCGCAACATTCGCGCAAGCGCGTCCCATCAGCATATA 780
Db 952 ATAGAAGACCATTTTGTATGTCACCTGCAAGATGAGCACCTATCTGCTGGCTTCATCAT 1011
Qy 781 TC 782
Db 1012 TC 1013

RESULT 10

US-09-452-239-1/c
; Sequence 1, Application US/09452239
; Patent No. 6465229
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Fader, Gary M.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase
; FILE REFERENCE: B01284 US NA
; CURRENT APPLICATION NUMBER: US/09/452,239
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 60/110,594
; EARLIER FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 50

SOFTWARE: Microsoft Office 97

SEQ ID NO 1
LENGTH: 891
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (806)
FEATURE:
NAME/KEY: unsure
LOCATION: (810)
US-09-452-239-1

Query Match 3.3%; Score 37; DB 4; Length 891;
Best Local Similarity 50.9%; Pred. No. 0.39;
Matches 88; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 363 GAAGACGCTCCGGTTGAAGAGCCAAAGGCTTATCTTCGTTACGTAGCGACGGGA 422
DB 648 GGAGCGCTCGGTAGCTAGGTAGTTGCTTGTGCGCGTCCACGAAGCGAAGTCA 589
QY 423 ACCAAGCGCGCTTATCGAAGATGTGCGCGCGCTTCTTGATGAAACCGCGCATCA 482
DB 588 AGCAGCGGTCTGCTCTTGTCCGCCACGAGTCTCCAGACGGGAGCGGGGCCCT 529
QY 483 CGGCCAACAATCGCTCGTGGCTCGAATCGGCGCTTGGGTGCGGAAACGGCT 535
DB 528 CGCGAAGTCGATCTTGTGGCCACGCGCGCTTGTGATGAGGAAGGCT 476

RESULT 11

US-09-252-991A-14103
Sequence 14103, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 14103
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14103

Query Match 3.2%; Score 36.4; DB 4; Length 975;
Best Local Similarity 45.9%; Pred. No. 0.61;
Matches 124; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 171 CGTGGCGGCAAGGCTATCGATCCAGGACAGGTCCTGTCATCGCGGTGAACGCTTTC 230
DB 168 CGCGGCTCGTCCGAGGATGTTCTTCGAGCGGTGCGCGCGGGTTCCTCCGATCC 227
QY 231 CGACCGGTGAGTGTATTCACGAACACCATATCTGCTGCTATCAGGTGACATA 290
DB 228 CGCGGTGATCGTCCGTACGCGGAGGCGGACGACCCGAACTCGAACATCCGCTCCCTG 287
QY 291 TCGCTGGAGCATGGCGGCTCGTATATCTTCACAAACCCGCGAGTGTGACATCGG 350
DB 288 CGCGGCATCTTCGAGTACTGTTGGAACCTTGGCGCGGACTCGATGTCCAGGCCA 347
QY 351 CGCGCGGCTTTGAAGACGTTCCCGGTGAAGAGCCAAAGGTCTTATCTTCGTTACGT 410
DB 348 TCTGCGCGAGCTGAGTCCGCGTGTGCTTGTGTCAGCGCGCTGCTGCGCGAGGTGCAACT 407
QY 411 AGGCGAGCGGCAACCAACCGCTTATCGA 440

DB 408 GGGCAGCGCGCGGACGCTGCGCGACGA 437

RESULT 12

US-09-252-991A-14168/c

Sequence 14168, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 14168
LENGTH: 1869
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14168

Query Match 3.2%; Score 36.4; DB 4; Length 1869;
Best Local Similarity 45.9%; Pred. No. 0.82;
Matches 124; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 171 CGTGGCGGCAAGGCTATCGATCCAGGACAGGTCCTGTCATCGCGGTGAACGCTTTC 230
DB 1720 CGCGGCTCGTCCGAGGATGTTCTTCGAGCGGTGCGCGCGGGTTCCTCCGATCC 1661
QY 231 CGACCGGTGAGTGTATTCACGAACACCATATCTGCTGCTATCAGGTGACATA 290
DB 1660 CGCGGTGATCGTCCGTACGCGGAGGCGGACGACCCGAACTCGAACATCCGCTGCGCTG 1601
QY 291 TCGCTGGAGCATGGCGGCTCGTATATCTTCACAAACCCGCGAGTGTGACATCGG 350
DB 1600 CGCGGCGCATCTCGAGTACTGTTGGAACCTTGGCGCGGACTGCGATGTCCAGGCCA 1541
QY 351 CGCGCGGCTTTGAAGACGTTCCCGGTGAAGAGGCAAGGTCTTATCTTCGTTACGT 410
DB 1540 TCTGCGCGAGCTGAGTCCGCGTGTGCTGTCAGCGCGCTGCTGCGGAGGTGCAACT 1481
QY 411 AGGCGAGCGGCAACCAACCGCTTATCGA 440
DB 1480 GGGCAGCGCGCGGACGCTGCGCGACGA 1451

RESULT 13

US-09-380-420C-1
Sequence 1, Application US/09380420C
Patent No. 6300544

GENERAL INFORMATION:
APPLICANT: Halkier, Barbara

Bak, Soren
Kahn, Rachel
Moller, Birger

TITLE OF INVENTION: Cytochrome P450 Monooxygenases
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,420C
FILING DATE: 12-No. 6300544-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-21251A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1929 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: P450ox
FEATURE:
NAME/KEY: CDS
LOCATION: 81..1673
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-420C-1

Query Match 3.2%; Score 36.4; DB 4; Length 1929;
Best Local Similarity 45.5%; Pred. No. 0.83;
Matches 130; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
QY 97 CGCTCGGAGCGAAATCGCGAAGGTACTCCCGTCCCGTCAAGCTCTCTGGGAGAGAG 156
DB 753 CGCTTCAGCACGCTGTGGAGCGCCATGACATGATGCGCAGCTTCTCCGCCGAGGAC 812
QY 157 ATTCTGCTCAATCGCTGGGCGCAAGGTCTATGCGATCCAGGACAGGTGCTGCAATCGC 216
DB 813 TTCTTCCCAACGCGCCGCGCTCGCGACCGGCTTCTCGGGCTTCTCGCCCGCGC 872
QY 217 GGTGTAAAGCTTTCCGACCGGGTCGAGTGCTATTCCAAAGAACCATATCTCTGTGTAT 276
DB 873 GAGCGCATCTTCAACGAGCTCGACGCTTCTTCGAGAAGGTCTATCGACCAGCATGGAC 932
QY 277 CAGGCTGGACATATCGCTGGGAGATGGCGCGCTCGTGCATATCTTCAAAACCCCGC 336
DB 933 CCGGCGCGCCCGTGGCGGACACAGCGGGGACCTCTGTCAGCTCTCTATCAACCTGTGC 992
QY 337 AGTGTGAGATCGGCGCGCGCTTTGAAGACGCTTCCCGTTGAAG 382
DB 993 AAGGAGCAGCAGCGACGCTCCGCTTCCAGGAGCCACGTCAGG 1038

RESULT 14
US-09-899-642A-1
Sequence 1, Application US/09899642A
Patent No. 6649814
GENERAL INFORMATION:
APPLICANT: Halkier, Barbara
Kahn, Rachel
Moller, Birger

TITLE OF INVENTION: Cytochrome P450 Monooxygenases
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,642A
FILING DATE: 05-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/380,420
FILING DATE: 12-No. 6649814-1999
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-21251A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1929 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: P450ox
FEATURE:
NAME/KEY: CDS
LOCATION: 81..1673
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-899-642A-1

Query Match 3.2%; Score 36.4; DB 4; Length 1929;
Best Local Similarity 45.5%; Pred. No. 0.83; 156; Indels 0; Gaps 0;
Matches 130; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
QY 97 CGCTCGGAGCGAAATCGCGAAGGTACTCCCGTCCCGTCAAGCTCTCTGGGAGAGAG 156
DB 753 CGCTTCAGCACGCTGTGGAGCGCCATGACATGATGCGCAGCTTCTCCGCCGAGGAC 812
QY 157 ATTCTGCTCAATCGCTGGGCGCAAGGTCTATGCGATCCAGGACAGGTGCTGCAATCGC 216
DB 813 TTCTTCCCAACGCGCCGCGCTCGCGACCGGCTTCTCGGGCTTCTCGCCCGCGC 872
QY 217 GGTGTAAAGCTTTCCGACCGGGTCGAGTGCTATTCCAAAGAACCATATCTCTGTGTAT 276
DB 873 GAGCGCATCTTCAACGAGCTCGACGCTTCTTCGAGAAGGTCTATCGACCAGCATGGAC 932
QY 277 CAGGCTGGACATATCGCTGGGAGATGGCGCGCTCGTGCATATCTTCAAAACCCCGC 336
DB 933 CCGGCGCGCCCGTGGCGGACACAGCGGGGACCTCTGTCAGCTCTCTATCAACCTGTGC 992
QY 337 AGTGTGAGATCGGCGCGCGCTTTGAAGACGCTTCCCGTTGAAG 382
DB 993 AAGGAGCAGCAGCGACGCTCCGCTTCCAGGAGCCACGTCAGG 1038

RESULT 15
US-09-252-991A-16243
Sequence 16243, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16243
LENGTH: 1314
TYPE: DNA

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16243

Query Match 3.2%; Score 36.2; DB 4; Length 1314;
Best Local Similarity 49.2%; Pred. No. 0.8;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY	745	GGCAACATTGGCAGCAAGCGCTCGCCATCAGCATATCGATCTGGTGGCGGCGTACTC	804
DB	812	GGCGAGATCTTACCAAGGTCGACGACAGCTATATGGGCGTGACCTCAGCGCATGATG	871
QY	805	AAGGTGGAACCGTGGCGCGATCCGAGCTCAGCGAGTTGCAATGTTACGTCCCGTCCGAT	864
DB	872	CTCGCGCGCGCCAGCGCGCTGGCGGAGCGCTGGAGCTGACGCGCTGGTGACCGCGAG	931
QY	865	GAGACCAAGCCACCTCTACTTCCAGACGCTGGGCAAGTGTGACGCTCAAGGAAGGGCA	924
DB	932	GGATCTCCAGGTGTCCAGCCAGACCTTGGGACCTCTCGGTGATCGACCGGGTTACC	991
QY	925	GACTCCTTCGAGC	937
DB	992	GACACCTGTGC	1004

Search completed: June 19, 2004, 07:22:48
Job time : 142.708 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 06:08:35 ; Search time 537.885 Seconds
(without alignments)
9693.657 Million cell updates/sec

Title: US-10-658-691-1

Perfect score: 1137
Sequence: 1 gtgcttaaccaacatcaat.....gcattcaggaagcagttga 1137

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1137	100.0	1137	17	US-10-658-691-1
2	46.2	4.1	3851	13	US-10-282-122A-28788
3	46.2	4.1	5496	13	US-10-282-122A-26551
4	43.2	3.8	536	17	US-10-338-110-119
5	43.2	3.8	498	16	US-10-259-194A-593
6	40.6	3.6	2823	13	US-10-282-122A-26243
7	40.6	3.6	2826	13	US-10-282-122A-28428
8	40.4	3.6	1801	13	US-10-425-114-20746
9	40.2	3.5	1830	13	US-10-282-122A-25435
10	40	3.5	1185	16	US-10-369-493-31790
11	39.8	3.5	606	12	US-09-758-759-130
12	39.8	3.5	109519	12	US-09-758-759-1
13	39.2	3.4	1308	13	US-10-282-122A-26553
14	39.2	3.4	2703	13	US-10-282-122A-28597

C 15	39	3.4	536	17	US-10-338-110-119	Sequence 119, Appl
C 16	38.6	3.4	690	9	US-09-738-626-2891	Sequence 2891, Ap
C 17	38.6	3.4	813	13	US-10-627-476-471	Sequence 471, App
C 18	38.6	3.4	3309400	9	US-09-738-626-1	Sequence 1, Appli
C 19	38.4	3.4	786	15	US-10-156-761-3595	Sequence 3595, Ap
C 20	38.4	3.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
C 21	38	3.3	594	13	US-10-142-426-10	Sequence 10, Appl
C 22	38	3.3	594	15	US-10-123-155-10	Sequence 10, Appl
C 23	38	3.3	594	15	US-10-146-731-10	Sequence 10, Appl
C 24	38	3.3	594	15	US-10-140-472-10	Sequence 10, Appl
C 25	38	3.3	594	15	US-10-141-761-10	Sequence 10, Appl
C 26	38	3.3	594	15	US-10-142-885-10	Sequence 10, Appl
C 27	38	3.3	594	15	US-10-158-790-10	Sequence 10, Appl
C 28	38	3.3	594	16	US-10-137-871-10	Sequence 10, Appl
C 29	38	3.3	594	16	US-10-140-923-10	Sequence 10, Appl
C 30	38	3.3	594	16	US-10-141-756-10	Sequence 10, Appl
C 31	38	3.3	594	16	US-10-141-758-10	Sequence 10, Appl
C 32	38	3.3	594	16	US-10-140-805-10	Sequence 10, Appl
C 33	38	3.3	594	16	US-10-140-864-10	Sequence 10, Appl
C 34	38	3.3	840	13	US-10-282-122A-33717	Sequence 33717, A
C 35	37.6	3.3	1581	13	US-10-369-493-27041	Sequence 27041, A
C 36	37.6	3.3	636	15	US-10-184-644-22	Sequence 15206, A
C 37	37.4	3.3	636	15	US-10-184-634-22	Sequence 22, Appl
C 38	37.4	3.3	919	15	US-10-184-644-258	Sequence 258, App
C 39	37.4	3.3	919	15	US-10-184-634-258	Sequence 258, App
C 40	37.4	3.3	919	15	US-10-063-685-700	Sequence 700, Appl
C 41	37.4	3.3	2838	9	US-09-954-456-700	Sequence 966, App
C 42	37.4	3.3	2838	9	US-09-954-456-966	Sequence 1143, Ap
C 43	37.4	3.3	2838	9	US-09-954-456-1143	Sequence 48, Appl
C 44	37.4	3.3	2838	15	US-10-175-523-48	
C 45	37.4	3.3	2838	15	US-10-175-523-48	

ALIGNMENTS

RESULT 1

US-10-658-691-1
; Sequence 1, Application US/10658691
; Publication No. US20040110258A1
; GENERAL INFORMATION:
; APPLICANT: Kayser, Kevin J.
; APPLICANT: Kilbane, John J.
; TITLE OF INVENTION: Method for Metabolizing Carbazole in Petroleum
; FILE REFERENCE: GTI-1512
; CURRENT APPLICATION NUMBER: US/10/658,691
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: US 60/409,562
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Spingomonas sp.
US-10-658-691-1

Query Match	100.0%;	Score 1137;	DB 17;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
Qy	1	GTGGCTAACCAACCATCAATCGCGAGCGGAGACCAAGGTTGGGAGCCTTATATCCGT	60	
Db	1	GTGGCTAACCAACCATCAATCGCGAGCGGAGACCAAGGTTGGGAGCCTTATATCCGT	60	
Qy	61	GCAGAACTCGGGTTCCGAAACCATTTGGTATCCGCTCGCGAGCGAAATCGCCGAA	120	
Db	61	GCAGAACTCGGGTTCCGAAACCATTTGGTATCCGCTCGCGAGCGAAATCGCCGAA	120	
Qy	121	GGTACTCCGTTCCGTCAGAGCTCCTGGGAGAGAGATTCTGCTCAATCGCGTGGGGGGC	180	
Db	121	GGTACTCCGTTCCGTCAGAGCTCCTGGGAGAGAGATTCTGCTCAATCGCGTGGGGGGC	180	

181	QY	AA	GGTCTATGCGATCCAGGACAGGTGCTGCATCGCGGTAAAGCTTTCGACCGGGTC	240
181	Db	AA	GGTCTATGCGATCCAGGACAGGTGCTGCATCGCGGTAAAGCTTTCGACCGGGTC	240
241	QY	GAGTGTATTCCAGAAACACATATCTGTGTGTTATCACGGCTGGACATATCGTGGGAC	300	
241	Db	GAGTGTATTCCAGAAACACATATCTGTGTGTTATCACGGCTGGACATATCGTGGGAC	300	
301	QY	GATGGCGGCTCTGTCGATATCTCTACAAAACCCCGGAGTGTGCAGATCGCGCGCGCGCT	360	
301	Db	GATGGCGGCTCTGTCGATATCTCTACAAAACCCCGGAGTGTGCAGATCGCGCGCGCGCT	360	
361	QY	TTGAAGACGTTCCCGGTTGAAGAGGCCAAAGGTCCTATCTTCGTTTACGTAGCGACGGC	420	
361	Db	TTGAAGACGTTCCCGGTTGAAGAGGCCAAAGGTCCTATCTTCGTTTACGTAGCGACGGC	420	
421	QY	GAACCAACGGCGTTATCGAAGATGTCCCGCCCGGCTTCCTTGATGAAAACCGGCCATT	480	
421	Db	GAACCAACGGCGTTATCGAAGATGTCCCGCCCGGCTTCCTTGATGAAAACCGGCCATT	480	
481	QY	CACGGCCAAACATCGGCTCTGGGCTCGAACTGGCGCTTGGTGGGAAAAACGGCTTGAT	540	
481	Db	CACGGCCAAACATCGGCTCTGGGCTCGAACTGGCGCTTGGTGGGAAAAACGGCTTGAT	540	
541	QY	CGCGGGCACGTTTCATTCACAGAAATTCGATCTCTGTGAAGGCCAACATATCATCTG	600	
541	Db	CGCGGGCACGTTTCATTCACAGAAATTCGATCTCTGTGAAGGCCAACATATCATCTG	600	
601	QY	CCGCTTCGCTTTGGCGTGGCGATCCCGACACAGCTTACGGCTTCCGAGGTTGTCGGGGC	660	
601	Db	CCGCTTCGCTTTGGCGTGGCGATCCCGACACAGCTTACGGCTTCCGAGGTTGTCGGGGC	660	
661	QY	AAGCCCAAAGGTGTTACGATCTGCTTGGCGAGCATTCGGTGC CGGTTTCGAAGGCATG	720	
661	Db	AAGCCCAAAGGTGTTACGATCTGCTTGGCGAGCATTCGGTGC CGGTTTCGAAGGCATG	720	
721	QY	ATCGAAGGCAAACTGCAATCCATGGCAACATTTGGCAGCAGCGCTCGCCATCAGCAT	780	
721	Db	ATCGAAGGCAAACTGCAATCCATGGCAACATTTGGCAGCAGCGCTCGCCATCAGCAT	780	
781	QY	TCGATCTGGCTGCCGGCGTACTCAAGGTGCAACCTGGCGCGATCCCGAGCTCACGCAG	840	
781	Db	TCGATCTGGCTGCCGGCGTACTCAAGGTGCAACCTGGCGCGATCCCGAGCTCACGCAG	840	
841	QY	TTCCGAATGATACGTCGGTTCGATGACACAGCCACCTCTACTTCCAGACGCTGGGCAAA	900	
841	Db	TTCCGAATGATACGTCGGTTCGATGACACAGCCACCTCTACTTCCAGACGCTGGGCAAA	900	
901	QY	GTCGTGACGTCAAGGAAGCGGCAGACTCTCTTCAGCGAGAAATTCACAGAAAATGGGTA	960	
901	Db	GTCGTGACGTCAAGGAAGCGGCAGACTCTCTTCAGCGAGAAATTCACAGAAAATGGGTA	960	
961	QY	GGCCTCCGGCTTAAACGGCTTCAATGATGACGACATCATGGGACGTGAATCGATGGAGCCG	1020	
961	Db	GGCCTCCGGCTTAAACGGCTTCAATGATGACGACATCATGGGACGTGAATCGATGGAGCCG	1020	
1021	QY	TTCTACGCTGATGATCGCGTGTGTCGGAAGAAATCTCTTTCAGCGCGGACCCGCGAATC	1080	
1021	Db	TTCTACGCTGATGATCGCGTGTGTCGGAAGAAATCTCTTTCAGCGCGGACCCGCGAATC	1080	
1081	QY	ATCGAGTGGCGGGGCTTGCAGTCACGACAAATCGCGGCATTCAGGAAGCACGTTGA	1137	
1081	Db	ATCGAGTGGCGGGGCTTGCAGTCACGACAAATCGCGGCATTCAGGAAGCACGTTGA	1137	

```

1  APPLICANT: Haselbeck, Robert
2  APPLICANT: Ohlsen, Kari
3  APPLICANT: Zyskind, Judith
4  APPLICANT: Wall, Daniel
5  APPLICANT: Trawick, John
6  APPLICANT: Carr, Grant
7  APPLICANT: Yamamoto, Robert
8  APPLICANT: Forsyth, R.
9  APPLICANT: Xu, H.
10 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
11 FILE REFERENCE: ELITRA.034A
12 CURRENT APPLICATION NUMBER: US/10/282,122A
13 CURRENT FILING DATE: 2003-02-20
14 PRIOR APPLICATION NUMBER: 60/191,078
15 PRIOR FILING DATE: 2000-03-21
16 PRIOR APPLICATION NUMBER: 60/206,848
17 PRIOR FILING DATE: 2000-05-23
18 PRIOR APPLICATION NUMBER: 60/207,727
19 PRIOR FILING DATE: 2000-05-26
20 PRIOR APPLICATION NUMBER: 60/230,335
21 PRIOR FILING DATE: 2000-09-06
22 PRIOR APPLICATION NUMBER: 60/230,347
23 PRIOR FILING DATE: 2000-09-09
24 PRIOR APPLICATION NUMBER: 60/242,578
25 PRIOR FILING DATE: 2000-10-23
26 PRIOR APPLICATION NUMBER: 60/253,625
27 PRIOR FILING DATE: 2000-11-27
28 PRIOR APPLICATION NUMBER: 60/257,931
29 PRIOR FILING DATE: 2000-12-22
30 PRIOR APPLICATION NUMBER: 60/267,636
31 PRIOR FILING DATE: 2001-02-09
32 PRIOR APPLICATION NUMBER: 60/269,308
33 PRIOR FILING DATE: 2001-02-16
34 Remaining Prior Application data removed - See File Wrapper or PALM.
35 NUMBER OF SEQ ID NOS: 78614
36 SOFTWARE: Patentin version 3.1
37 SEQ ID NO 28788
38 LENGTH: 3851
39 TYPE: DNA
40 ORGANISM: Mycobacterium tuberculosis
41 US-10-282-122A-28788

```

```

RESULT 2
US-10-282-122A-28788
; Sequence 28788, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
;

```

Sequence 26551, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26551
TYPE: DNA
ORGANISM: Mycobacterium bovis
US-10-282-122A-26551

Query Match 4.1%; Score 46.2; DB 13; Length 5496;
Best Local Similarity 46.9%; Pred. No. 0.00033;
Matches 144; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY	614	CGCCTGGGATCCCGACAGCTTACGCGTTCGAGGTTGCTGCGGGCAAGCCCAAGGTG	673
Db	2648	CGCTGGAAGAAGCCGCGAGGTACGGTGCCCTGTTGATGCGTGACCCCGCGGG	2707
QY	674	TTTACGATCGTTGCGAGCATTCGGTCCGGTTTTGAAGCGATGATCGAAGGCAAC	733
Db	2708	GTGACCCGCTGTGCGCGGCTGCAGTCGTCGTCGCGACCGCATCCGACCTCGAGC	2767
QY	734	CTGCAATCATCGACATTCGCAAGCGGTGCGCATCAGCATATCGATCTGCTGTC	793
Db	2768	TGGCATCCGCGGCGGACGAAGCAGCTGCGCGGCGGATCAGCGAATCTCTCGGAGC	2827
QY	794	CGGGCGTACTCAAGTCGAACCGGTGGCCGGATCCGAGCTCAGCAGTTTCAATGTCAG	853
Db	2828	GGATCGTTGTTGTTCTGTCGCCGCCGAGCAACGACGATCGATCGCGAGCGGTCAAC	2887
QY	854	TGCGGTCGATGAGACCAACCGACCTTACTTTCAGACGCTGGCAAGTCGTGACGTCAA	913
Db	2888	TGAGTTGGACGCAACACGACGCTGCTGTCGCCAGCGTGTGACGACCGTGACCGCAA	2947
QY	914	AGGAAGC 920	

Db 2948 TGGGTGC 2954
RESULT 4
US-10-338-110-119
Sequence 119, Application US/10338110
Publication No. US20040023254A1
GENERAL INFORMATION:
APPLICANT: Fuhrmann, Jeffrey J.
APPLICANT: Romesser, James A.
TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
FILE REFERENCE: HER-0056
CURRENT APPLICATION NUMBER: US/10/338,110
CURRENT FILING DATE: 2003-01-07
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.2
SEQ ID NO 119
LENGTH: 536
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus Sequence
FEATURE:
NAME/KEY: misc feature
LOCATION: (7)..(7)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (21)..(21)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (69)..(69)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (87)..(87)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (213)..(213)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (226)..(226)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (261)..(261)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (273)..(273)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (347)..(347)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (357)..(357)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (398)..(398)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:

```
; NAME/KEY: misc feature
; LOCATION: (419)..(419)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (509)..(509)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (517)..(517)
; OTHER INFORMATION: n is a, c, g, or t
US-10-338-110-119

Query Match      3.8%; Score 43.6; DB 17; Length 536;
Best Local Similarity 13.2%; Pred. No. 0.00096;
Matches 67; Conservative 230; Mismatches 208; Indels 2; Gaps 1;

Qy 373 CCGGTTGAAGAGCCAAAGTCTTATCTTCGTTTACGTAGGCGAGCGGCAACCAAGCCG 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 SVRHKTVMVMGVMVWVBREHNSYSRHKTSKSGNCRCTWYCGGCRYSRVSWVWVYTSRGR 60

Qy 433 CTTATCGAAGATGTGCGCGCGGCTTCCTTGATGAAAACCGCGCCATTACGCGCCACAT 492
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 MRRCVTGNGTGSRYBYCVWYSKMNWVVDMMBSBKWGVARYKSGAYCARTWCGAYWY 120

Qy 493 CGGCTCGTGCCTCGAACTGCGCTTGGGTGCGGAAACCGCTTTGATGCGGGGCGACGTC 552
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 SMSAHRCSKCTACVTSBTSBGSHKVRGMVSVVVBDBHATSKYGGHTGYGCSGSGTG 180

Qy 553 TTCAATCACAAGATTCGATCTGCTGGAAGGGCAACGATATCATCTTCGCGCTTGCTTT 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 HTDCCBACSAVBDSSCBWYCTGCTBRARKVNSTVTVYRSSBMYNSKCRBVSRYVM 240

Qy 613 GCGCTGGCATCCGACACAGCTTAGCGCTTCGAGGTTGCTCGGGGCAAGCCCAAGGT 672
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 MYRCCSMVSVATCCVMSRNTTGGARHTKWCNWTWCRFCRMCRCRHSCSSMYKGG 300

Qy 673 GTTACATCTGCTGCGAGCATCTGGTGC--CGGTTTTCGAAGGATGATCGAAGGCA 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 CSRTGMVKWRTYTSKSSDVNRBCSTSVARWSMGSDDSYVWSBSSGNGCVKRBVNTSR 360

Qy 731 AACCTGCAATCCATGGAACATGGCAGCAAGCGCGTGGCATCAGCATATCGATCTGGC 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 TSGSVRSYSHCRBBSRSTAGWVCGSYKWTWVDNMKNSANCGBRVTMDSEHSANC 420

Qy 791 TGCGGGCGGCTACCAAGTGAACCGTGGCGCGGATCCCGAGCTCAGCGATTCGAATGT 850
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 GVSYVGBYCGSYBMSVWVHVRBRBVRDMHRTVSTSGCSTSYTGGAYCRYBADCRMD 480

Qy 851 ACGTGGCGGTCGATGACACGACCCACC 877
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 SCVYVRMSVHBYKYWSKRYRSMHYC 507

RESULT 5
US-10-259-194A-593
; Sequence 593, Application US/10259194A
; Publication No. US2004010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghassemian, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP

; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NAME/KEY: misc feature
; LOCATION: (509)..(509)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (517)..(517)
; OTHER INFORMATION: n is a, c, g, or t
US-10-338-110-119

Query Match      3.8%; Score 43.2; DB 16; Length 498;
Best Local Similarity 46.6%; Pred. No. 0.0013;
Matches 138; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 698 CGGTGCGGTTTTTCGAAGGCATGATCGAAGCAAACTGCAATCCATGCGCAACATTGGCA 757
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 CGHAGGCGGGTCCACCGCCAGCAGCGCGGAGCTGGCCCGCTACGGGACGGGACG 259

Qy 758 GCAAGCGCGTCCCATCAGCATATCGATCTGGTTCGCGGGCGTACTCAAGTCCGAACCGT 817
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 GCGACGCGCGGCTCAGCGTGGACTACGCTCGCTCGCCGACATCGACGCTTCTTCCAGT 319

Qy 818 GCGCGGATCCCGAGCTCAGCGAGTTCGATCTGGTTCGCTGAGTGCAGTGCAGACCGCCACC 877
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 GCGCCAAAGTGCATGAGCAGCAGCCACTACTACTAGCAGGCTCATGACCCCTCGCGCG 379

Qy 878 TCTACTTCCAGACGCTGGGCAAGTCGTGAGCTCAAGAGGAGCGGCGGACTCTTTCGAGC 937
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 CTTCTTCGCGCGCGCCATCGGTTCGCGGCGAGCAGCAGCAGCGCGCCCGCAT 439

Qy 938 GAGATTCACGAAAATGGTAGGCTTCGCGCTTACGGCTTACGGCTTCAATGATGACGAC 993
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 GCTGGGAGGACGCGCAGCGCATTCGCTCTGGAGCTTCAGCTTCCCGCTCTGGAC 495

RESULT 6
US-10-282-122A-26243
; Sequence 26243, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28428
; LENGTH: 2826
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28428

Query Match          3.6%; Score 40.6; DB 13; Length 2826;
Best Local Similarity 45.8%; Pred.No. 0.019;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY      346 ATCGCCCGCGCGTGTTTGAAGACGTTCCCGGTTGAAGAGCGCCAAAGTCTTATCTTCGTGTT 405
      |||||
Db       667 AGCGCCGGATCATCCGACTGTCTGTCGCCCTGGTGCAACAGGCCACACCGTGGCGCACTG 726

QY      406 TACGTAGCGGACGGCGGAACCAACGCCGCTTATCGAAGATGTGCGCGCCCGCTTCCTTGAT 465
      |||||
Db       727 GTGCGCGTGTGCGCGCGACTTGTGTGGCCTCATCGCTGATCGCGCGCGCCCGAGAGATCGGC 786

QY      466 GAAAAACCGCGCATTCACGGCCAAACATCGGCTCTGTGCGCTTCGAACATGGCGCTTGGGTGCG 525
      |||||
Db       787 GCTGACGTGCGCTTTTGGCACCACACAACGGTTCGGAGTGCCGATGGGTTTGGCGGCCCG 846

QY      526 GAAAAACGGCTTTGATGCGGGCAGTC 552
      |||||
Db       847 CATGCCGGGTACCTTGGGTGTCAGCC 873

```

```

US/10-425-114-20746
; Sequence 20746, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20746
; LENGTH: 1801
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: LIB3245-118-C5_FLI
US-10-425-114-20746

Query Match          3.6%; Score 40.4; DB 13; Length 1801;
Best Local Similarity 51.1%; Pred. No. 0.019;
Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy      372  CCGGGTTCAAGAGGCCAAAGGCTTACTTCGTTTACGTAGGCGAGCGGAAACCAACGCC 431
Db      668  CCGGGAATGCGGCMACACAGGCCTCCGGCTTCTCGCCCGTGGCGACCAACAGAGCAGGCGC 609

Qy      432  GCTTATCGAAGATGTGCGCCCGCGCTTCCTTTGATGAAAAACGGGCCATTACGGGCCAACA 491
Db      608  CGAATGCTACTCATGTGCGCGCGCGCTCCGCGGTTTGGTGACGACGCGCGCAGAGCCCGACG 549

```



```
/
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(606)
; OTHER INFORMATION: evbu
US-09-758-759-130

Query Match      3.5%; Score 39.8; DB 12; Length 606;
Best Local Similarity 45.4%; Pred. No. 0.019;
Matches 143; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 772 ATCAGCATATCGATCTGGCTGCCGGCGTACTCAAGGTGCGAACCGTGGCCCGATCCCGAG 831
Db 10 ATCAGCTCGACCTCGGATGCGCGCCGCCCTCTGATCGAGCTGGCCCTCGCCGGCGG 69
Qy 832 CTCAGCAGTTCGAATGTAAGTCCGCGTTCGATGAGACGACCGACCTTACTTCCAGACG 891
Db 70 GTCCGCTAGTGTGGAACCTGTGTGTCAGGACCCCGGCCCGCCACCGGGAACCGATC 129
Qy 892 CTGGGCAAGTCTGACGTCAAGGAAGCGGAGACTCTCTTCGAGCGAGATTCACGAA 951
Db 130 CTGGAGAGGTCTCTGGCGGATGCGCGCGGACCGCGACACCCGCTCTCTGGGTG 189
Qy 952 AAATGGGTAGGCTTCGCGCTTAAACGGCTTCAATGATGACGACATCAGCGACGTGAATCG 1011
Db 190 CAGCGCTGCGCCAGCGCTGCGGACCGGATCTCTCGGCGACCTGGTGGCGCGGGCGTG 249
Qy 1012 ATGGAGCGCTTCTACGCTGATGATCGCGTGTGTCGAGAAATCCTGTTGAGCCGAC 1071
Db 250 GTGGCGAGCTCGAGGACCGAGCTGGCTTCATCTCAGCTGACCCGCTACCCGACGCG 309
Qy 1072 CGCGCAATCATCGAG 1086
Db 310 GAGCCACCGTGGAG 324

RESULT 12
US-09-758-759-1/c
; Sequence 1, Application US/09758759
; Publication No. US20040101832A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tim X.
; APPLICANT: Horan, Ann C.
; TITLE OF INVENTION: Evernimicin Biosynthetic Genes
; FILE REFERENCE: ID0983K US
; CURRENT APPLICATION NUMBER: US/09/758,759
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/175,751
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Micromonospora carbonacea
US-09-758-759-1

Query Match      3.5%; Score 39.8; DB 12; Length 109519;
Best Local Similarity 45.4%; Pred. No. 0.16;
Matches 143; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 772 ATCAGCATATCGATCTGGCTGCCGGCGTACTCAAGGTGCGAACCGTGGCCCGATCCCGAG 831
Db 83879 ATCAGCTCGACCTCGGATGCGCGCCGCCCTCTGATCGAGCTGGCCCTCGCCGGCGG 83820
Qy 832 CTCAGCAGTTCGAATGTAAGTCCGCGTTCGATGAGACGACCGACCTTACTTCCAGACG 891
Db 83819 GTCCGCTAGTGTGGAACCTGTGTGTCAGGACCCCGGCCCGCCACCGGGAACCGATC 83760
Qy 892 CTGGGCAAGTCTGACGTCAAGGAAGCGGAGACTCTCTTCGAGCGAGATTCACGAA 951
Db 83759 CTGGACGAGTCTTGGCGGATGCGCGCGGACACCGCGCAACCCCGCTCTCTGGGTG 83700

/
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(606)
; OTHER INFORMATION: evbu
US-10-658-691-1.rnpb

Query Match      3.4%; Score 39.2; DB 13; Length 1308;
Best Local Similarity 45.7%; Pred. No. 0.042;
Matches 137; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

Qy 641 GTTCCCGAGGTGCTGCGGGCAAGCCCAAGAGGTGTTTACGATCTGCTTGGCGACATTCCG 700
Db 737 GTACCGTCGAGGCGCTGGAAGAGCCCTGATCGGTATCCAGGTGGACGAGGTGGTGC 796
Qy 701 TGGCGGTTTTCGAGGCGATGATCGAAGGCAAACTGCAATCTTCATCGCAATTCGACGCA 760
```

Db 797 TCGGGGTGATCGACCGGGTGTGCGGGCATTAACGAAACCAACGTCACACCTGGCGTGG 856
Qy 761 AGCGGTGCGCATCAGCATATCGATCTGGTGTGCGGGGTGATCTCAAGGTGCAACCGTGGC 820
Db 857 CTTCCGATGCGGTGATCATCGGTTTCAATGTGCGCGCGCGAAGCAAGCGGACCGAGCTGG 916
Qy 821 CGGATCCGAGCTCAGCAGTTCGAATGGTACGTGCGGGTCCGATGAGACAGCACCTCT 880
Db 917 CCAGCCCGAAGGGTGGAGATCGCTACTACTCTACCGGGGATCGACGAGA 976
Qy 881 ACTTCCAGACGCTGGGCAAGTCTGAGCTCAAGGAAGCGGAGACTCTCTTCGAGCGAG 940
Db 977 TCGAGCAGCGCTGCTGCGTGGCTGTCTCAAGCGCATCTACGAGGAAACCCAGCTGGTGGTGG 1036

RESULT 14

US-10-282-122A-28597
; Sequence 28597, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28597
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28597

Query Match 3.4%; Score 39.2; DB 13; Length 2703;
Best Local Similarity 45.7%; Pred. No. 0.056;
Matches 137; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

Qy 641 GTTCCGAGTTGCTCGCGGCAAGCCAAAGGTGTTTACGATCTGCTGGCGGAGCATTCGG 700
Db 2129 GTACCGTCAGCGCTGGAAGAGGCCCTGATGGGTATCCAGGTGACGACGAGGTGGTGC 2188

Qy 701 TCGCGGTTTTTGAAGGATGATCGAAGGCAAACTGCAATCCATGGCAACATTGGCAGCA 760
Db 2189 TCGCGGTGATCGACCGGGTGTGCGGGCATTAACGAAACCAACGTCAACTGGCGTCGG 2248
Qy 761 AGCGGTGCGCATCAGCATATCGATCTGGTGTGCGGGGTGATCTCAAGGTGCAACCGTGGC 820
Db 2249 CTTCCGATGCGGTGATCATCGGTTTCAATGTGCGCGCGCGAAGCAAGCGGACCGAGCTGG 2308
Qy 821 CGGATCCGAGCTCAGCAGTTCGAATGGTACGTGCGGGTCCGATGAGACAGCACCTCT 880
Db 2309 CCAGCCCGAAGGGTGGAGATCGCTACTACTCTCGGTCACTACCGGGGATCGACGAGA 2368
Qy 881 ACTTCCAGACGCTGGGCAAGTCTGAGCTCAAGGAAGCGGAGACTCTCTTCGAGCGAG 940
Db 2369 TCGAGCAGCGCTGCTGCGGCTGTCTCAAGCGATCTACGAGGAAACCCAGCTGGTGGTGG 2428

RESULT 15

US-10-338-110-119/c
; Sequence 119, Application US/10338110
; Publication No. US20040023254A1
; GENERAL INFORMATION:
; APPLICANT: Fuhrmann, Jeffrey J.
; APPLICANT: Ronesser, James A.
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
; TITLE OF INVENTION: Communities
; FILE REFERENCE: HER-0056
; CURRENT APPLICATION NUMBER: US/10/338,110
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33)..(33)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (69)..(69)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (87)..(87)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (213)..(213)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (226)..(226)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (261)..(261)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (273)..(273)
; OTHER INFORMATION: n is a, c, g, or t

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 04:04:20 ; Search time 1331.47 Seconds
(without alignments)
10742.394 Million cell updates/sec

Title: US-10-658-691-2

Perfect score: 330

Sequence: 1 ataccgcaaggtccgcgt.....tgcggaagccacagtcattga 330

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_btgo_hum.*
40: em_btgo_mus.*
41: em_btgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	330	100.0	4154	1	AF442494	AF442494 Sphingomo
2	330	100.0	13651	1	AB095953	AB095953 Sphingomo
3	115.2	34.9	12376	1	AE014404	AE014404 Brucella
C 4	115.2	34.9	12429	1	AE009536	AE009536 Brucella
C 5	114.2	34.6	345897	1	AP002395	AP002395 Mesorhizo
C 6	114	34.5	300140	1	AP005956	AP005956 Bradyrhiz
C 7	107	32.4	1351	1	CCFDBX	XS1607 C.crescentu
C 8	107	32.4	10158	1	AE006011	AE006011 Caulobact
C 9	106.4	32.2	4151	1	RCFDXE	Y11304 R.capsulatu
C 10	105.6	32.0	6407	1	AB121977	AB121977 Xanthobac
C 11	100.6	30.5	349746	1	BX572605	BX572605 Rhodosphe
C 12	89	27.0	309350	1	AP005215	AP005215 Corynebac
C 13	86.2	26.1	10495	1	AE008062	AE008062 Agrobacte
C 14	86.2	26.1	10883	1	AE009097	AE009097 Agrobacte
C 15	86.2	26.1	306250	1	SME591788	AL591788 Sinorhizo
C 16	85.8	26.0	299450	1	AP005938	AP005938 Bradyrhiz
C 17	85.2	25.8	6040	1	AB062679	AB062679 Bradyrhiz
C 18	84.4	25.6	4379	1	ASP311718	AJ311718 Acinetoba
C 19	81.2	24.6	347894	1	BX640431	BX640431 Bordetell
C 20	81.2	24.6	348842	1	BX640446	BX640446 Bordetell
C 21	79.6	24.1	348934	1	BX640417	BX640417 Bordetell
C 22	79.2	24.0	349497	1	BX640440	BX640440 Bordetell
C 23	74	22.4	4398	6	AX645690	AX645690 Sequence
C 24	74	22.4	4398	6	AX661050	AX661050 Sequence
C 25	74	22.4	8766	6	AX645680	AX645680 Sequence
C 26	74	22.4	8766	6	AX661040	AX661040 Sequence
C 27	74	22.4	12342	6	AX645679	AX645679 Sequence
C 28	74	22.4	12342	6	AX661039	AX661039 Sequence
C 29	74	22.4	23656	1	AF333761	AF333761 Rhodococc
C 30	74	22.4	23656	6	AX645678	AX645678 Sequence
C 31	74	22.4	23656	6	AX661038	AX661038 Sequence
C 32	73.6	22.3	321	6	AX661046	AX661046 Sequence
C 33	73.6	22.3	321	6	AX645686	AX645686 Sequence
C 34	72.8	22.1	6458	1	RSU17130	U17130 Rhodococcus
C 35	58	20.6	146584	2	AC079815	AC079815 Trypanoso
C 36	67.2	20.4	4416	1	AF096864	AF096864 Pseudomon
C 37	65.6	19.9	306	6	AX120694	AX120694 Sequence
C 38	65.6	19.9	306	6	BD162811	BD162811 Novel pol
C 39	65.6	19.9	318	6	AX120696	AX120696 Sequence
C 40	65.6	19.9	318	6	BD162813	BD162813 Novel pol
C 41	65.6	19.9	441	6	AX065549	AX065549 Sequence
C 42	65.6	19.9	328050	1	AP005275	AP005275 Corynebac
C 43	65.6	19.9	349880	6	AX127144	AX127144 Sequence
C 44	64.8	19.6	2497	3	DMHSPG3	X06542 Drosophila
C 45	64.4	19.5	322774	3	CEY73F8A	AL132862 Caenorhab

ALIGNMENTS

RESULT 1
AF442494
LOCUS
DEFINITION Sphingomonas sp. GTIN11 carbazole operon, partial sequence.
ACCESSION AF442494
VERSION AF442494.1 GI:17227030
KEYWORDS
SOURCE Sphingomonas sp. GTIN11
ORGANISM Sphingomonas sp. GTIN11
Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
Sphingomonadaceae; Sphingomonas.
REFERENCE 1 (bases 1 to 4154)
AUTHORS Kilbane,J.J. II, Daram,A., Abbasian,J. and Kayser,K.J.
TITLE Isolation and characterization of Sphingomonas sp. GTIN11 capable
of carbazole metabolism in petroleum

Pred. No. is the number of results predicted by chance to have a

JOURNAL Biochem. Biophys. Res. Commun. 297 (2), 242-248 (2002)
MEDLINE 2222130
PUBMED 12237109
REFERENCE 2 (bases 1 to 4154)
AUTHORS Kayser K.J., Daram A., Abassian, J. and Kilbane, J.J. III.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2001) Biotechnology, Gas Technology Institute,
1700 South Mount Prospect Road, Des Plaines, IL 60018, USA
FEATURES
source
1. .4154
/organism="Shingomonas sp. GTIN11"
/mol_type="genomic DNA"
/db_xref="taxon:174635"
-1. .4154
/gene="carbazole operon"
650. .1786
/gene="carAa"
650. .1786
/gene="carBa"
/codon_start=1
/transl_table=11
/product="carbazole dioxygenase"
/protein_id="AAL37976.1"
/db_xref="GI:17227031"
/translation="WANOPISTAEERTKWEPIYRAKLGFRNHWYPVRLASIBAGTPV
PVKLCEKILLNRVGKQVIAIQOCLHRGVTLSDREVCYKNTICSWYHGWTYRWDDG
RVLDILTPSGVOIGRRALKTPVEEAKGLIPVYGDGEPTLIEDVPPGLDENRAI
HGQRLVATNSGDAENGDPDAGHFIHKNILVKGNDIILPLGAPGDDOLTSEVA
ACKPGVYDLLSEHSPVPEGMIEGKPAHNGIGSKRVASISLWPLVPLKVEWPDP
ELTQEWYVFDVETSHLIFQTLGKVVTSKEADSFERFHEKWKWGLALNGFNDDIIMA
RESMPFFYADDDSEELFEPDRAIIEMRGLASQHRNGIOEAR"
1735. .2067
/gene="carBa"
1735. .2067
/gene="carBa"
/note="meta cleavage enzyme subunit"
/codon_start=1
/transl_table=11
/product="2-aminobiphenyl-2,3-diol 1,2-dioxygenase"
/protein_id="AAL37977.1"
/db_xref="GI:17227032"
/translation="WAGACQSAQSRHSGLTGMTTASYPVNRMLMOELTNPQNVLELFR
ADREALYERYGLSSAQRALDEGGFALTAVGLHPVLQMHHPMLNPNAPDFVSVKAY
RKMVDNRG"
2060. .2863
/gene="carBb"
2060. .2863
/gene="carBb"
/note="meta cleavage enzyme catalytic subunit"
/codon_start=1
/transl_table=11
/product="2-aminobiphenyl-2,3-diol 1,2-dioxygenase"
/protein_id="AAL37978.1"
/db_xref="GI:17227033"
/translation="WAEIVGVFATSHVLSFGSPDQAQLRVVVEGMEIGRRVRLLRPD
LTIIVGSDHLNITITLQPFVTVSDTFTPLGMDIEORPAGNRAFAESLCARAD
RFDLAGEBLRPHGVNPLMTDPDRIIPVPIVYNNMTPTPTAARAOLGDTVAE
AVGLDLPRLRVVAVATGLSWINIPGHGEVNAFEDSRVIAAFQSGDMRLRAIDTE
TLKXNAGSGLFVWVMAAATLPGRRAEKIYYPNPQMVTGNGGIAIV"
2906. .3730
/gene="carC"
2906. .3730
/gene="carC"
/codon_start=1
/transl_table=11
/product="meta cleavage compound hydrolase"
/protein_id="AAL37979.1"
/db_xref="GI:17227034"
/translation="WVLAATAGRSVTVRGMEITYHEQSGDVVVLVHGGAGDSMGN
WRGMPVLADEVRIAVDMGLGRTAKPADPVPFQOARTDHLAGFLDALGSLNVALV
GNSMGGLAVGVERPGLVRLNMGSLVSKIDPALEPVLGVDFTRGEMIRLVRA
LTTNDFQIDAMIDRYALSDVPETERAYSATQWIRDOGGLYEDDYIRITAPTLLI
VNGKLDKVLNAYKFLFELIGSMGYINPDGCHWMIHHPVDFPARTTAATFIAAQ"

gene 3771. .4100
/gene="carAc"
3771. .4100
/gene="carAc"
/codon_start=1
/transl_table=11
/product="ferridoxin"
/protein_id="AAL37980.1"
/db_xref="GI:17227035"
/translation="MTAKRVIVFRAAGGFELHVEYTAGVSLMEAAVINGVDGIEAVCG
GACACATCHVYVGPWMLDKPSETEDMLDCVAERAPHSRLSCQIRLTLDLDGLTL
ELPKAQS"
ORIGIN
Query Match 100.0%; Score 330; DB 1; Length 4154;
Best Local Similarity 100.0%; Pred. No. 9.3e-56;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACCGCAAAGGTCGCGGTGATCTTCGCGCAGCGCGGGCTTCGAGCATCTGTCGAA 60
Db 3771 ATGACCGCAAAGGTCGCGGTGATCTTCGCGCAGCGCGGGCTTCGAGCATCTGTCGAA 3830
Qy 61 ACCGAGCGGAGTATCGCTCATGGAAGCGCGCTTCTGAACGCGCTGACGATATCGAA 120
Db 3831 ACCGAGCGGAGTATCGCTCATGGAAGCGCGCTTCTGAACGCGCTGACGATATCGAA 3890
Qy 121 GCCTTTTCGCGGGCGCTGTGCTGCGCCACGTGCGCTTACGTTGCGCCCGAGTGG 180
Db 3891 GCCTTTTCGCGGGCGCTGTGCTGCGCCACGTGCGCTTACGTTGCGCCCGAGTGG 3950
Qy 181 CTAGATCGCTGAAAACCGCGAGTAGACCGAAGAAATGCTCGATTGCGTAGCGGAA 240
Db 3951 CTAGATCGCTGAAAACCGCGAGTAGACCGAAGAAATGCTCGATTGCGTAGCGGAA 4010
Qy 241 CGTGGCGCATTCGCGGCTGCTGCGCAGATCGGCTTACCGACCTGCTCGACGCGCTG 300
Db 4011 CGTGGCGCATTCGCGGCTGCTGCGCAGATCGGCTTACCGACCTGCTCGACGCGCTG 4070
Qy 301 ACCTGGAATGCGGAGGACAGTCATGA 330
Db 4071 ACCTGGAATGCGGAGGACAGTCATGA 4100
RESULT 2
AB095953
LOCUS AB095953 13651 bp DNA linear BCT 04-FEB-2003
DEFINITION Shingomonas sp. KAI car operons, complete cds.
ACCESSION AB095953
VERSION AB095953.1 GI:28201219
KEYWORDS
SOURCE Shingomonas sp. KAI
ORGANISM Shingomonas sp. KAI
Bacteria; Proteobacteria; Alphaproteobacteria; Shingomonadales;
Shingomonadaceae; Shingomonas.
REFERENCE 1
AUTHORS Inoue, K., Nojiri, H., Nakai, S., Endoh, T., Urata, M., Ashikawa, Y.,
Saiki, Y., Yoshida, T., Habe, H. and Omori, T.
TITLE Divergent structures of carbazole-degrading car operons isolated
from several gram-negative bacteria
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 13651)
AUTHORS Nojiri, H. and Inoue, K.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-2002) Hideaki Nojiri, Biotechnology Research
Center, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo
113-8657, Japan (E-mail:anojiri@mail.ecc.u-tokyo.ac.jp,
Tel:81-3-5841-3067, Fax:81-3-5841-8030)
FEATURES
source
1. .13651
/organism="Shingomonas sp. KAI"
/mol_type="genomic DNA"
/strain="KAI"
/db_xref="taxon:164608"

```
gene complement (127. .2199)
CDS /gene="ORF10"
complement (127. .2199)
/gene="ORF10"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56754.1"
/db_xref="GI:28201220"
/translation="MISEKIGPOHLERKAILYVQSSVHQVHLNRRSSLOVAMRGRL
TANGSHIETIDDLGSAAGVTRAGFDWVAEVLGRVGAAREVSRFARNSRDW
QQIEMCRVDTVLIDQETVYAPQGNDRLLGLKGLSNEYELDLRQSLRSLK
RREGVVAAPGVGVGADRIEKDPDRRQVQAIALVDFKVAELGSAQALLMTIEHGLD
LPKRNQGVVVRPSTVATHTMIENPIYGGAYAGKSAATGYDGTAMRSRKRAR
DEWLALPGSHGVGVSERSEDIKMVDNVPVSOHGAPKGGDLAGLIRCRRCGR
KLVTYTGAKHDIPRYSCRGLLDNGSPRCIAFGGLRVDDAIERALLQVLEPGAIAS
VEAEOADRDQDVQVLMEDLEAARYADRAEROVDAADPNELVAELEFVFWAL
TRAGEVARIVADASTALPALPGLSKIDGLASLDAQVWNPASQDARLKRIVRTLIQ
EVVADHDHASEIVLLHWGVLHTDLRPRRKGQNSADSIIAAVRLNLIANDD
LIAGLNRNLVGTGNGRWRTRVTRALSRHRIPVFRTVADGPAPWLNLQSAARHTGV
ASKTLRIAAAGBIKGIHLPPEGWIFCRTELDGSAAHHLAKARQNPKYPTGSHPDQ
QTLFSSTT"
gene 2502. .2852
CDS /gene="ORF9"
2502. .2852
/gene="ORF9"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56755.1"
/db_xref="GI:28201221"
/translation="MGRTHPCQCARSRVEHPRNLEPPVRSRSLATAQINRTGFD
LMDVDMPASGPMFAENFENFGFVGASPCCTTTSARTAVLAIGHRQKRHRIRP
FVPLRSTSVIRMR"
gene 3162. .3785
CDS /gene="ORF8"
3162. .3785
/gene="ORF8"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56756.1"
/db_xref="GI:28201222"
/translation="MKYNAAERFHEYGITKSITPNVFIHTLOEVNDGLAGRY
FMDVADGIALQSDSDVHTPKTPKDALNKAERYFDEWPPSAKNRIAKFDVG
QGYLRDAADFQATERLYHCVLVCTFTYPHVNTLGFURSQANLDRILMYVPEDN
RKQRVMEKLQKAYVKARYSKHYRISEELTWLGEQVEELGRVYHAV"
gene 3862. .4188
CDS /gene="ORF7"
3862. .4188
/gene="ORF7"
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56757.1"
/db_xref="GI:28201223"
/translation="MQRDSGDSRQNGRSATKNRDILLAAFPVSCANAGFRKREAGEGS
NFGFGQREIIDIYAEVSDCILDVGNAEQDLYGAELACGFVNQRRLRSRVRATFP
TVETNR"
gene complement (4409. .5089)
CDS /gene="carR"
complement (4409. .5089)
/gene="carR"
/codon_start=1
/transl_table=11
/product="transcriptional regulator of car operon"
/protein_id="BAC56758.1"
/db_xref="GI:28201224"
/translation="MISTTWTSLDKESPAHRAAAFTIRGAILEARFAPQRLVESELS
SLTGLGRGPVREARLLAAEGLVAIERNKGAIVARASKAMMVEFVEVRELLGLEAAR
AALRAATSPHLESTERALDEESLENDDSGIILMEANEHLHLQVLADEASVIERVLV
OMQLEPSSIFFRLMSPTRHSREDHIGILNTILKGDGEAAEQMRHARVTRTAALFA
DLPDELLR"
gene 5192. .6328
CDS /gene="carAa"
5192. .6328
/gene="carAa"
/codon_start=1
/transl_table=11
/product="terminal oxygenase component of CAR
1,9a-dioxygenase"
/protein_id="BAC56759.1"
/db_xref="GI:28201225"
/translation="MANQPSIAERTKWKVEPIYIAKLGFRNHWYPVLASIAEGTPV
PVKILGKILINRVGGKVYAIQDRCLARGVTLSDRVCYSKNTISCVYHGWYRWDDG
RLDILITVPGSVQIRRALKTFPVVEAKGLIFVVVGCEPTLIEDVPPGFLDENRAI
HGOHLVASNRELGAENGFDAGHVIHKNSILVKNNDIILPLGAPGDDOLTRSEVA
AKQKGVYDILLGERSVPVPEGNIEGKPAHKNISGRVAISISILWLPGLVKEWPPDP
ELTQFQWYVDDETSHLYFQTLGKVTSKEAADSFEFFHEKFWGLALNGFNDDDIMA
RESKEPFYADDRGWSSEILFEPDRAIIEMRGLASQHNRIQEAR"
gene 6328. .6609
CDS /gene="carBa"
6328. .6609
/gene="carBa"
/codon_start=1
/transl_table=11
/product="subunit of meta cleavage enzyme"
/protein_id="BAC56760.1"
/db_xref="GI:28201226"
/translation="MGTTASYPVNRLMQELFTNPQNVLPFRADREALYVYGLSSAQR
AALDEGGGALTAVGLHFLVLMQHHFMLTNPAFVSVKAYRKWDRNG"
gene 6602. .7405
CDS /gene="carBb"
6602. .7405
/gene="carBb"
/codon_start=1
/transl_table=11
/product="subunit of meta cleavage enzymes"
/protein_id="BAC56761.1"
/db_xref="GI:28201227"
/translation="MAEIVGVFAFTHVLFGSDGDAQALRVVEGNEIGRRVRLLRPD
LLIVIGSDHFNITRLOPFTVGVSDTFTFLGDMDEIQRFPAGNAPFAELCARAAD
RFDLAQGSELPRDGVVPLFIDPGRIPVPIYVYNNMTPTTAAARAAQLGIVAE
AVGLDPSHLRVVVVATGTLGSHWINIPGHGEVNAEFORRVIAAFQSGDMRWLRADTE
TLLKNAGNGGLEIVNWMAATLFGRAEKIYBEMPQWMTGMGIAIV"
gene 7448. .8272
CDS /gene="carC"
7448. .8272
/gene="carC"
/codon_start=1
/transl_table=11
/product="meta cleavage compound hydrolase"
/protein_id="BAC56762.1"
/db_xref="GI:28201228"
/translation="MVLAAAGRSVTVRGMEITYHEQSGDVVVLHGGGAGASMGNG
WRGVPLADRYVIAVDMLGFGRTAKPADFPVFSQAARTDHLGFLDALGSLNVALV
GNSMGASALGVAVRPGVRLKVLMSGAGLSKIDPALPEVLGVYDFTFREGMIRLVA
LTTNQQDIDAMIDYRVALSVDPTRRAYSATMOWIROQGLYVEDDYIRITPTLI
VNGKLDKVPPLANAYKPLELIGPSWGYIMPCGHWAMIEHPVDPARTAAFEAAQ"
gene 8313. .8642
CDS /gene="carAc"
8313. .8642
/gene="carAc"
/codon_start=1
/transl_table=11
/product="ferredoxin component of CAR 1,9a-dioxygenase"
/protein_id="BAC56763.1"
/db_xref="GI:28201229"
/translation="MTAKRVIFRAAGGFHELVETEAGVSLMEAAVLNGVVGISAVCG
GACACATCHVYVGPWEWLDALKEPPSEDEMLDCVAERAPHSRSLSCQIRLTLDDGLTL
ELPAQS"
gene 8684. .10999
```

```

CDS
/gene="ORF11"
8684..10999
/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56764.1"
/db_xref="GI:28201230"
/translation="WCKLLMSALAVLASAPAFQAQTDGVEVDASELEIITVTABR
VSENLQDSFVAISAGEAENIRAINVSDIGKFTPLSFGSFAAASGSSNVTFVLR
GIGQDFNMVDPGVIYLDGVYISRSVGALETRDVEISQVLRGPQGTGFKGTIGG
AVITSRPDEQLGNIEATIGYINRDIIVGARINVPVDSFVYSGIALTRDGYER
LVDGQKQNSYSGRFAEMDAPDFNLLSADYTRRERAVATLLDVNENAIFPL
FNFVFNAPGLPPSPSCYNACMLTGDKTNSGPNKRSVSDVNGVSNASWDA
GPVTIKSTAYRELDASAFVDIDQSLAIQRTENDYTONFOSQFQLGDCGFGKLYN
TLGLYHHEGDELDINSLFPAIDPDSGSGVNDYVAAEQEFERATONLNLTLGRVT
SDKRFSDQVLRVHPAAQALPFGPFTPPFLNPNVTSMAAGSPQLRVGDIYMPAV
QRTTANETPVTLDYKFTPDMGLFTFSFGSGGTQKRVFPPEPAPAFEPFAR
VLEIGVVELFRRLNSALDTRYISDLQIIVNDGFPKVRNAGKARIRFEGEQL
VYDAPRLTGAVYLDABEYVDARAAPVTVDSALANTPKWTANAGITVEYDGEAGR
VVLQGDWSVRSATYKDAINTPSLQKQGYSVFNMSATFDAAQGSVTGGITNLTKRYI
ISGSDLNFTIGSTGYVSRPPEWYKLGVSF"
/complement(11435..12808)
/gene="ORF12"
complement(11435..12808)
CDS
Query Match 100.0%; Score 330; DB 1; Length 13651;
Best Local Similarity 100.0%; Pred. No. 8.3e-56;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCGCAAGTCCGCGTGATCTCCGCGAGCGCGCGGCTTCGAGCATCTGTCGAA 60
DB 8313 ATGACCGCAAGTCCGCGTGATCTTCGCGAGCGCGCGGCTTCGAGCATCTGTCGAA 8372
QY 61 ACCGAAGCGGAGTATCGTCTATGGAAGCGCGGCTTCGAAAGCGGTCGACGGTATCGAA 120
DB 8373 ACCGAAGCGGAGTATCGTCTATGGAAGCGCGGCTTCGAAAGCGGTCGACGGTATCGAA 8432
QY 121 GCGGTTTGGCGGGCGCGCTGTGCTGCGCACGTGCGCACGTTTACGTTGCGCCCGAGTGG 180
DB 8433 GCGGTTTGGCGGGCGCGCTGTGCTGCGCACGTGCGCACGTTTACGTTGCGCCCGAGTGG 8492
QY 181 CTAGATCGCGTGAACCGCGGAGTGACGGAAGCGGCAATGCTCGATTGCGTACGGAA 240
DB 8493 CTAGATCGCGTGAACCGCGGAGTGACGGAAGCGGCAATGCTCGATTGCGTACGGAA 8552
QY 241 CGTGC CGCGGATTCGCGGCTGCTCCGAGATCCGCGCTTACCGACCTGCTCGACGCGCTG 300
DB 8553 CGTGC CGCGGATTCGCGGCTGCTCCGAGATCCGCGCTTACCGACCTGCTCGACGCGCTG 8612
QY 301 ACCCTGGAACCTGCGGAAGGCACAGTCATGA 330
DB 8613 ACCCTGGAACCTGCGGAAGGCACAGTCATGA 8642

RESULT 3
LOCUS AE014404 12376 bp DNA linear BCT 30-SEP-2002
DEFINITION Brucella suis 1330 chromosome I section 90 of 190 of the complete genome.
ACCESSION AE014404
VERSION AE014404.1 GI:23347845
KEYWORDS
SOURCE
ORGANISM Brucella suis 1330
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Brucellaceae; Brucella.
1 (bases 1 to 12376)
REFERENCE
AUTHORS Paulsen,I., Seshadri,R., Nelson,K.E., Eisen,J.A., Heidelberg,J.F.,
Read,T.D., Dodson,R.J., Umayam,L.A., Brinkac,L.M., Beanan,M.J.,
Daugherty,S.C., Deboy,R., Durkin,A.S., Kolonay,J.F., Madupu,R.,
Nelson,W.C., Ayodeji,B., Kraul,M., Shetty,J., Malek,J.A., Van
Salzberg,S.E., Riedmuller,S., Tettelin,H., Lindler,L., Halling,S.M.,
Boyle,S.M.,
and Fraser,C.M.
The Brucella suis genome reveals fundamental similarities between
animal and plant pathogens and symbionts
Proc. Natl. Acad. Sci. U.S.A. 99, 13148-13153 (2002)
2 (bases 1 to 12376)
Paulsen,I., Seshadri,R., Nelson,K.E., Eisen,J.A., Heidelberg,J.F.,
Read,T.D., Dodson,R.J., Umayam,L.A., Brinkac,L.M., Beanan,M.J.,
Daugherty,S.C., Deboy,R., Durkin,A.S., Kolonay,J.F., Madupu,R.,
Nelson,W.C., Ayodeji,B., Kraul,M., Shetty,J., Malek,J.A., Van
Salzberg,S.E., Riedmuller,S., Tettelin,H., Lindler,L., Halling,S.M.,
Boyle,S.M.,
and Fraser,C.M.
Direct Submission
Submitted (14-AUG-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..12376
/organism="Brucella suis 1330"
/mol_type="genomic DNA"
/strain="1330"
/db_xref="taxon:204722"
/chromosome="I"
/complement(66..4739)
/locus_tag="BR1024"
complement(66..4739)
/locus_tag="BR1024"
/notes="similar to GP:15074750; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN29946.1"
/db_xref="GI:23347846"
/translation="MASKGAQKLDLEVKALEQALDIDFGDDIDMDLGFADDEGFS
IDLEQISAAEELVAVQAAEPVNEPVVKAASVITTAATVPATVPRTVPANDDSRD
AASAPAIARTEKPAVVAEPVSAQTAAPKAAVATATPVATVPRTVPANDDSRD
KATAAPRRITEHSKLYMTTVAISVLMAGGVAISKAINEPVESLOATKDFPTAPA
GLGVAGVAPLPIAFWPGFQQLVRAQMNAASMTDAKLLQLEPAAGDVRVSTLGO
AVREVAAMNIEIETLVALETLVQSEVNRKAYSDNEVRIKSLVSDLGNEREA
VSHAKRVSSISGHEQREELSSASNIIRNVLSAQSLGALLDGERLISLINE
SGAIADAEKRTGDIETTSGEAFNLITRIATLDEQSTIGSKUTALDERTT
GIANVLSATQTVSFDFRLANLENTLSDRGLSLAEFAHLDNSTEKLNALE
TRSQINENLIARTBIAETFSGRNSLGMVDIDKVKIDDLTISBSVGNVLI0KA
AAEGLAESRDLAATLEGETERSVIRGHADTLAARTGDIQNAIDASASALNEVT
DRAAIAERTSALQAMSNVSLDASADHARSLERTASLHNVIAQNAQMLAQM
DERAAMNENRNRVLAFTFERIGSLQSLAENQSALERLIDRANMAESLGG
RDAPFALGHEVETVEERTSGLOAVLNHNDLAHVLDGHEQTETRAAEIRSLIDS
TASLSQIDHGEILEQRTSLQNAIANSAAISSAFEGTGTIERETMEKALSIG
VDNVRALEQAGVAGKREKIGEAATLAAEAKAGALDGFBSFARLIDNLSIG
TEARLGERDAIAGRLGDIESRLTDELGTIARIAADTAARTSETSLASHSESFSVAD
NLAGTETRLVERADATRLGDIISRITWELGSVEARIQAVTDTTAVTLSEERELNA
VLAARSOEITKILNDTAEPLOVRLADSGRLAQOLEATHAATDRLSNAALVNLA
SRTAETIAVQAKVGLSDNVSLIDRLAASNGELGLIDAATNLADIDGLVDSTT
SPVENTNRAAMFQASTGLDLSNIGTJRLSLDNTLSQIADIADIRFHEGKVLSSADM
INSAQNLVSTLEERHQAQKGLSEKSEGVKQMSFEALVAFAGRAQGRVRS
AEKRVSSDIVQAAKFSAADEIRTAASEIRAEIDTFRGKGLVDMPABAKET
TAMRKAVSQINAKELAEIVNKGRLVDGGRADRPVSPAPAPVAPVAPAP
IAPLACTDVAVRSPVEAPNOKAPAAAEKPRGWVSDLLARASREEEAAQKQVQPA
ATPSSHVVEVSLNSLSDIARAIDHEASVELWDRYRGRNVTFRRLYTLKQQTDFD
DIRKQYTDGCFRVDYRMDQRLLEDVARNDRDNMTYTQTLTSDTGKVTMLAHA
SGRLR"
4954..5118
/locus_tag="BR1025"
4954..5118
/locus_tag="BR1025"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN29947.1"
/db_xref="GI:23347847"

```


CDS	<pre> /gene="folk" /locus_tag="BR1031" 8313..8846 /gene="folk" /locus_tag="BR1031" /note="similar to GP:15074743; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="2-amino-4-hydroxy-6- hydroxymethylidihydropteridine pyrophosphokinase" /protein_id="RAN29952.1" /db_xref="GI:23347852" /translation="MSTPPKRLAWLGLGNIIDPIISMASALRMLDGSDDTRVVSVS PVYTPPKGKTDQAFNACQAIETAQPIELIAACLDVRSLSKRVLRWGWGPIIDI DILVQDETGRAITWQEPALFLPHRHERAFVLVPLNDIAPSLPVGARTVAEWAADC DRAEMKARTDQGWLE" complement(8961..9518) /locus_tag="BR1032" complement(8961..9518) /locus_tag="BR1032" /note="similar to GP:14021684; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="RAN29953.1" /db_xref="GI:23347853" /translation="MDSSTELIMKNCIAACILLPVLSVSPVMAADKDKQFFOTIEG QWSGGEIVAGYKGTXYCNLAGTTPDPTVGTGLDCTCRVGVFSOPMKATVTRVGDG YAKFNDAAGDLGVTSKGVNKNVVFGLNRKQLNGAMLARVSDENTMNTVTSVRVE KELVPVIGMSLRVDITIAVGSIAKN" complement(9770..10843) /locus_tag="BR1033" complement(9770..10843) /locus_tag="BR1033" /note="identified by match to TIGR protein family HMM TIGR01620" /codon_start=1 </pre>
	<pre> Query Match 34.9%; Score 115.2; DB 1; Length 12376; Best Local Similarity 60.6%; P-Id. No. 3.1e-13; Matches 189; Conservative 0; Mismatches 123; Indels 0; Gaps 0; </pre>
Qy	19 GTGATCTTCGCGACGCGGGCGGTTCGAGCATCTGTCTCGAAACGGAACGGGGAGTATCG 78
Db	5366 GTATTCTGATCCCGCGATGGCGCCACGGAACCGAGGTCGAGGCGGACACGCGTCCAGC 5425
Qy	79 CTGATGGAACGCGCGCTTCGAAACGCGTGGACGCTATCGAAGCGCTTTGCGGGGGCGCC 138
Db	5426 GTTATGGAACGCGCATTCGCAACGGCATTCGCGTATTGATGCGGATCGGTGGCGCT 5485
Qy	139 TGTGCTTCGCCACGTCGCCACGTTTACGTTGTCGCCCGCGAGTGGCTAGATCGCTGAACCG 198
Db	5486 TGGCGCTTCGCAACCTCGCCATGCTCTATGTCGATGACGACTGGGCGGATACGGTTCGGCGGG 5545
Qy	199 CCGAGTCAGACCAAGACGAATGCTCGATTTCGCTAGCGGAGCGTGGCGCCGATTCGCGG 258
Db	5546 CCGGATCCATGGAAGAGGATATGCTGGATTTCGCTATGAATCCCGCCGACCTCGCGC 5605
Qy	259 CTGTCCTGCCAGATCCGCCCTTACCGACCTGCTCGACGCGCTGACCGCTGGAAGCTCCGAG 318
Db	5606 CTTTCCTGCCAGATCCCGGTTACCGACGATCTGGAGGGGCTGTGCTTCAGTGCCTCCGAA 5665
Qy	319 GCACAGTCATGA 330
Db	5666 CGCCAGAACTGA 5677

RESULT 4

AE009536/c

LOCUS

DEFINITION

Brucella melitensis 16M chromosome I, section 93 of 195 of the

linear

BCT 20-MAR-2003

complete sequence.	/transl_table=11
AE009536.1	/product="2-AMINO-4-HYDROXY-6-HYDROXYMETHYLDIHYDROPTERIDINE PYROPHOSPHOKINASE"
AE009536.1	/protein_id="AAL52135.1"
Brucella melitensis 16M	/db_xref="GI:17982911"
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Brucellaceae; Brucella.	/translation="MSTPKPLRAMLGICGNIDDDPIISASALRMLDGRSDTRVSVS
1 (bases 1 to 12429)	PVTRTPGKGTQAMPHNACQIETALQLELIAVCLDVERSLKRVRERKRGRIIDI
Los T., Ivanova, N., Anderson, I., Bhattacharyya, A., Lykidis, A.,	DLVMDQETGRAITMQEPALFPHPRHERAFVLAFLNDIAPSLPVGARTVAEWAADC
Reznik, G., Jablonski, L., Larsen, N., D'Souza, M., Bernal, A.,	DRAMEKARTDAGWLE"
Mazur, M., Goltsman, E., Selkov, E., Elzer, P. H., Hagius, S.,	complement (1315. .1674)
O'Callaghan, D., Letesson, J.-J., Haselkorn, R., Kyrpides, N. and	/gene="BMEI0955"
Overbeek, R.	/genes="BMEI0955"
The genome sequence of the facultative intracellular pathogen	/EC number="4.1.2.25"
Brucella melitensis	/codon_start=1
Proc. Natl. Acad. Sci. U.S.A. 99 (1), 443-448 (2002)	/transl_table=11
11756688	/product="DIHYDRONEOPTERIN ALDOLASE"
2 (bases 1 to 12429)	/protein_id="AAL52136.1"
DelVecchio, V.G., Redkar, R.J., Patra, G. and Mujer, C.	/db_xref="GI:17982912"
Direct Submission	/translation="MYTIRIMCAPFAHGVDFDEHKLCQRFYDAVLVDVDEGNSLES
Submitted (13-NOV-2001) Institute of Molecular Biology and	PNVSGVDYVEVTVPYQ"
Medicine, University of Scranton, Scranton, PA 18510, USA	complement (1693. .2532)
3 (bases 1 to 12429)	/gene="BMEI0956"
Elzer, P.H. and Hagius, S.	complement (1693. .2532)
Direct Submission	/gene="BMEI0956"
Submitted (13-NOV-2001) Department of Veterinary Science, LSU Ag	/EC number="2.5.1.15"
Center, 111 Dairymple Building, Baton Rouge, LA 70803, USA	/codon_start=1
4 (bases 1 to 12429)	/transl_table=11
Kaputra, V., Los, T., Ivanova, N., Anderson, I., Bhattacharyya, A.,	/product="DIHYDROPTEROATE SYNTHASE"
Lykidis, A., Reznik, G., Jablonski, L., Larsen, N., D'Souza, M.,	/protein_id="AAL52137.1"
Bernal, A., Mazur, M., Goltsman, E., Selkov, E., Haselkorn, R.,	/db_xref="GI:17982913"
Kyrpides, N. and Overbeek, R.	/translation="MTKQWLAHERSLHGKTAIMGVLNVTDPDSFGDGRYNDLDR
Direct Submission	IAAETWLEGATIVDVGSFPPGATADTAQTEASRVVPIRELVRPDCIISVDTY
Submitted (13-NOV-2001) Integrated Genomics, Inc., 2201 W. Campbell	RASTARLAPAGAHVNDVWGLQREPEIQALETAGLVIMTSDRLNTPDVIDEQ
Park Drive, IL 60612, USA	FMFLNSLKTAGDAGIEPISVLDVPGFGKEDIEDIALMASELQAFGLPVLVGTIS
5 (bases 1 to 12429)	RKEFVGAMTQADPRNRDGTATSVLRADGADI FRVHNVAFNRDALAVADAILQSN
Letesson, J.-J.	VRI"
Direct Submission	2729. .3310
Submitted (13-NOV-2001) Unite de Recherche en Biologie Moleculaire,	/gene="BMEI0957"
Laboratoire d'Immunologie et de Microbiologie, Universite of Namur,	2729. .3310
61 rue de Bruxelles, Namur 5000, Belgium	/genes="BMEI0957"
6 (bases 1 to 12429)	/codon_start=1
O'Callaghan, D.	/transl_table=11
Direct Submission	/product="hypothetical protein"
Submitted (13-NOV-2001) Faculte de Medecine, INSERM U431, Avenue	/protein_id="AAL52138.1"
Kennedy, Nimes 30900, France	/db_xref="GI:17982914"
Location/Qualifiers	/translation="MAILSPQAEAAALRKFQYVYTINGKTAADLKDLSRSGPFLKKT
1. .12429	QHPHGAERIRFPAKYRGPCKQVQDVYVNVAKYSLPRWKQRKASPELAIIM
/organism="Brucella melitensis 16M"	DTLLQDIRHEESHIVARSHASEMERELRSLRADCASLRADIDKVTSRVMEADE
/mol_type="genomic DNA"	AQQYDLVETINFENRPERLLTYRLRMRQTH"
/strain="16M"	complement (3315. .3656)
/db_xref="taxon:224914"	/gene="BMEI0958"
/chromosome="I"	complement (3315. .3656)
200. .670	/gene="BMEI0958"
/gene="BMEI0953"	/EC number="1.6.4.5"
200. .670	/codon_start=1
/genes="BMEI0953"	/transl_table=11
/codon_start=1	/product="THIOREDOXIN REDUCTASE"
/transl_table=11	/protein_id="AAL52139.1"
/product="Hypothetical Protein"	/db_xref="GI:17982915"
/protein_id="AAL52134.1"	/translation="MAENLDMAGORLIGTNRGTTTPADAIIVIASGLGCFNGGQIVR
/db_xref="GI:17982910"	PDPTRWGLRERCNGNAIVDPETATSCPGVFAIGDACHYFGKLLILSGFHEALMTQ
TVGMTLDGTCRVGFSGPMKATVTRVGDGVAGKNDGKGLDVTSGVNGKVVFG	AIRYIYIAQOQ"
LNKQKQNGAMLARVSDPTNNVTVSVRVEKLPVIGMSLKRVDITAVGSIANK"	complement (3956. .4297)
complement (785. .1318)	/gene="BMEI0959"
/gene="BMEI0954"	complement (3956. .4297)
complement (785. .1318)	/gene="BMEI0959"
/genes="BMEI0954"	/codon_start=1
/EC number="2.7.6.3"	/transl_table=11
/codon_start=1	/product="FERREDOXIN, 2FE-2S"
	/protein_id="AAL52140.1"
	/db_xref="GI:17982916"

```
/translation="MSOGIKMTKIVFVSADGATRTVEADSGSSVMEEAIRNGIPGID
AECGACACATCHYVDDDDWADTVGGDPMEEDMLDPAYEVPTSLSCQIRVTDDE
GLVQVPERQN"
/gene="BMEI0960"
/complement(4398) .4679)
/gene="BMEI0960"
/complement(4398) .4679)
/codon_start=1
/transl_table=11
/product="Hypothetical Protein"
/protein_id="AAU52141.1"
/db_xref="GI:17982917"
/translation="MWSKVSLSFSPGHAFQRLTIKRVKREYLQKXHLQDAVADQDS
PVFFISSYKIIWRFSPCGKVLAPAFADFDAGAWMSRGAAYGG"
/gene="BMEI0961"
/587. 9567
/gene="BMEI0961"
/587. 9567
/gene="BMEI0961"
/codon_start=1
/transl_table=11
/product="Kinesin-like protein"
/protein_id="AAU52142.1"
/db_xref="GI:17982918"
/translation="MFWGFAQLVKRAQEMONARSMTDAALKLLOPEAVAGRVSTLG
QAVREVAANNEGIERLARAVELETLVQSEVQLERAYSQDNVRSLSVSDLGNERE
AVVBAEKVRSISGAHEQREELSSNIIKRNVLSSQSLGALLGDSGERLISIN
ESGALADAIEKRTDIGNRTTGEAPANLDPRIATLDEQRTSGKLTQALDERT
TGIANVLASDTWVSFDTLANLENTLSDRGSLABEFARHALDNSTELNAL
EATROQINENLIARLTAIEGERSNGLAMVDIKVIGDITISGVGNVLDK
AAEFGKLSRDLAATLEGETERSVVGADTIAARTGDIQNAIDASALANEV
TDHAAIILTERTSALOQMSNVSLDASPADHARSLEERTSLHNVLNAGNLAQL
MDERAAAMRENFRNVRILARTFDERIGSLQSLAENQSLERILDERARNAESLGA
GRDEFALHGEKRVISERTSGLOQVNNNDHALVHDGHEQTIETRAAEIRSLTLD
STASLTLDHGLBILORTSLQNAIANSAAISSAFEGGTGIIERTETMBKALSI
GVNVRALQSQAVAKREKIGEAATLAAAKAGALDGFESFARLIDNLS
GTEARLGERDAIAGRLDIESRLTDELGTIEARIADTAARTSETLASHSESFSAVA
DLNAGTRELIVRAADATRLDIDTIGRTIWEGLSVIARIQVTDITAVTLEERTREL
AVLAARSOEILKIINDAEPLVOELADSLGSLAQELQEAHTAATDRLSERNAALNAL
ASRTAEITIAVQQAQKVSNDVSELDRSLAASGELGLIDAIATRLNADIDGRVNST
TSFVNTNRAQMFQASTGLDLSNIGLITLSTLNTLSQIADIRDFDEHKGVLSSAD
MINSQANGLVSTVEERQALDKLASGLIVERSEKVGKMQSFEALVASAFORAGQTRV
SAEKORRESVDIVBQAQKFSATDEIRTAETRIEALDITRGLKRGVLMFAEAKE
TTTAMRKAVSQINAKELABIVNKGSLVDVGEGRADRPVSPAPASRPAPVQAPVRA
PIALADTVAVRPVPEPNOKPAPASAEKPRGVSDLLARASEREEBAPKQVP
AATPRSPHVVESLNSLVDIARAIDHEASVELMDRVYRGERNVFRRLVTLKGOQTF
DDIKRKVQDGEFRVAVDRVNDPQRLLEDVANDRDNWVTQIYILSDTGKVTMLAH
ASGRIR"
/complement(9807) .10364)
/gene="BMEI0962"
/complement(9807) .10364)
/gene="BMEI0962"
/codon_start=1

Query Match 34.9%; Score 115.2; DB 1; Length 12429;
Best Local Similarity 60.6%; Pred. No. 3.1e-13;
Matches 189; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Qy 19 GTGATCTTCCCGCAGCGCGGCTTCGAGCATCTGTCGAAACCGAGCGGAGTATCG 78
Db 4267 GTATTCGTATCCGCGATGGCGCCACGACGACCGAGTTCAGGCGGACGCGTCCAGC 4208
Qy 79 CTCATGGAAGCGGCGGCTTGAACGGCGGTGACCGGATCGAAGCCGTTTCGGGGCGCC 138
Db 4207 GTTATGGAAGCGGCGCATTCGACCGGATTCGCGGATTCGCGGATTCGCGTGGCGCT 4148
Qy 139 TGTGCTGCGCCAGTGCAGTTCGTTAGTGGCCCCGAGTGGCTAGATGCGCTGAACCG 198
Db 4147 TGCCTCTCGCAACTGCGCATGTCTATGTCATGACGACGTCGGCCGATACGGTCGGCGG 4088
Qy 199 CCGAGTGAAGCCGAGACGAATGCTCGATTGCGTAGCGGAACGTCGCGCGCATTCGCGG 258
Db 4087 CCGGATCCAAATGGAAGAGGATATGCTGATTCGCTTATGAAGTCCGCCCGACCTCGCGC 4028
```

```
Qy 259 CTGCTCTCCAGATCCGCTTACCGACCTGTCGACGCGCTGACCTGGAACCTGCCGAAG 318
Db 4027 CTTTCTCTCCAGATCCGCTTACCGACCTGTCGACGCGCTTGTCTTACGTCGCCGAA 3968

Qy 319 GCACAGTCATGA 330
Db 3967 CGCCAGAACTGA 3956

RESULT 5
AP002995/c 346897 bp DNA linear BCT 15-MAY-2001
LOCUS Mesorhizobium loti DNA, complete genome, section 2/21.
DEFINITION AP002995 BA000012
ACCESSION AP002995.2 GI:14021442
VERSION 1
KEYWORDS Mesorhizobium loti
SOURCE Mesorhizobium loti
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Phyllobacteriaceae; Mesorhizobium.
REFERENCE 1 (sites)
AUTHORS Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,
Watanabe,A., Idegawa,K., Ishikawa,A., Kawashima,K., Kimura,T.,
Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,
Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpou,S., Sugimoto,M.,
Takeuchi,C., Yamada,M. and Tabata,S.
Complete genome structure of the nitrogen-fixing symbiotic
bacterium Mesorhizobium loti
JOURNAL DNA Res. 7 (6), 331-338 (2000)
MEDLINE 21082930
PUBMED 11214968
REFERENCE 2 (bases 1 to 346897)
AUTHORS Kaneko,T.
Direct Submission
Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp, rhizobase/,
URL:http://www.kazusa.or.jp/rhizobase/,
Tel:81-438-52-3935 (ex.2338), Fax:81-438-52-3934)
On May 11, 2001 this sequence version replaced gi:11994963.
COMMENT Location/Qualifiers
source
/organism="Mesorhizobium loti"
/mol_type="genomic DNA"
/strain="MAFF303099"
/db_xref="taxon:381"
74. .160
/note="trns-TGA"
/complement(240) .1109)
/gene="msl18587"
/complement(240) .1109)
/gene="msl18587"
/note="unknown protein"
/codon_start=1
/transl_table=11
/db_xref="GI:14021443"
/db_xref="BAB48055.1"
AGNVLIVSNKXVLYGKWSLSEDFALLDNQKRVGPGANTRLTKNQLPISEHPDPTV
DUALPIQPIVEMRGQKGIIFNLGRTIAPDNQIALRAASVLMVGFSPGIMDE
KNLPLVRRGILAVNYLTDYRGNTNFVVDIAFGSSGSPVFAFFENMLPDGSGVTM
LAQPVYFLVYLHSGFSTATGKIVPVPVPTSDQIAQTQMLHLGYCVKASRVESLVE
TIRAKISSEPTASG"
/complement(1118) .1657)
/gene="ml10476"
/complement(1118) .1657)
/gene="ml10476"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48056.1"
/db_xref="GI:14021444"
```

/translation="MFPISALIGFEKQGWGALAGEAMAQTAEDNFRLEVWDRE
KNLVETICSPSPSTSOAAWQAIIRRRPGMVLHYNSRWWEKIIITPCEPKVPQTV
DSVHAGQDVALGDLREWHKLRWCKNCKSHAEVPAALIKRYGEAALFSTVERALFC
TCDRGPPVRLHKLPRN"
2193..2558
/gene="mlr0478"
/gene="mlr0478"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48057.1"
/db_xref="GI:14021445"

gene

CDS

/translation="NALAGGALMKWRLYFVVMTLAVTSGCAADYLNNDYTVTLASGD
ANHNTSLQTVDPFNPNSNNTHEGQGRIVGVQVRSPISAGGYSGNCPTNSVA
ADGSHCKESAESRPGAM"
2651..3412
/gene="mlr0479"
2651..3412
/gene="mlr0479"
/codon_start=1
/transl_table=11
/product="succinoglycan biosynthesis protein; ExoI"
/protein_id="BAB48058.1"
/db_xref="GI:14021446"

gene

CDS

/translation="MFLRLVTVAAASALATQFVNHNGGSLPEINLONLKKPZVAI
SGVASVIGDTEMHGQVRVFNIGIDAPESAQYCDDAKPEYPCGRRSAEALDGLAAS
KPVQCAFVTDWRYGREGVGDCTADGSSVAAVMVHGQALDMPKYGNGAYAAQAKAKA
AKQGLWGFQFQAPWRAQSDDAQTPPTAFLATGNPGCNKIGNISASGERIVHLP
GQKYAVTVISQKGRWFCSEAEAAVAGWRSSK"
3429..3782
/gene="mlr0480"
3429..3782
/gene="mlr0480"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48059.1"
/db_xref="GI:14021447"

gene

CDS

/translation="MKPHYRAAAVALSHSSGRNVSSVYSELKYNNIDASVOGVN
VRYDYNNSCHIDGTLPSTLYHGQNSHIDLQPSQGYDGYDGNSSHFLTVSQNSA
ELYDGEQGFQFST"
complement(5516..5947)
/gene="mlr0481"
complement(5516..5947)
/gene="mlr0481"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48060.1"
/db_xref="GI:14021448"

gene

CDS

/translation="MNSDRTGDPSTGTGFLPHGQFKKVGSRSRVSRQADTANEHA
ATAMTVLGMGLAGDIRAAEILRLRYGADPFLPKNRVLPDVENPTSENLAVALSAA
WSGEISPDKAKOLLDAIAKGKGIIEADALLRVEQLAVKR"
complement(6065..6883)
/gene="mlr0482"
complement(6065..6883)
/gene="mlr0482"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48061.1"
/db_xref="GI:14021449"

gene

CDS

/translation="MIKDFKIDGASSRGLFQWNVNGSVISAFWPNYSTWSPDIDA
ASALNNGVLAAGNAVIGAAERLKRTEIRDVSLAIAFYAAPFRYLDRDAIDERN
VIGSRNSLAQIPIPNTPFSIAADSLERQEVRRMLDQDQGEAFDWMQSPGTVL
NAVAPSIDLTIPSKDQQLAGLIVVEEPAIKWARTINDPSNDNLATAANPLKNKMSDDL
NRLAKARDGLKLRQAVGSDAKELLAVIAPASQAGNLTPEAEFWMGMQKVA"
complement(7133..7753)
/gene="mlr0483"
complement(7133..7753)
/gene="mlr0483"
/note="unknown protein"

gene

CDS

/codon_start=1
/transl_table=11
/db_xref="GI:14021450"
/translation="MANYILRLTVNMKNRASTITADHANPAKLGOIADHAFWLMSGT
REWASVPNFEQYRDECDMDVSGYEGNIRITTFDKTYQSYSFNDAATTDGHPFV
WDTLGCADNMVVSSTHFSGIINGVWSASVDQSDNWQFEDSSSTVSEMIDRLHLPW
RDPFTGVAROSKELPPAIRALVKGVRSKDRMTAPNQFAENIH"
complement(7827..10169)
/gene="mlr0485"
complement(7827..10169)
/gene="mlr0485"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48063.1"
/db_xref="GI:14021451"
/translation="MRDGSVRLPPELLPYRQVVRVRLBQKSPKPAKPKYASNGR
ADTNPATNSDYPTAVALASQYAGIFVFAENDDFFVIDLDGCDATIGELSAEAA
AITRMPFAAWELSGEGEHIIFGRCDKVALGFRNKWAGCWFYTKRFVAFPGGI
QGNIDIDWTNLAILVPARGVGEQDASQEAVERDPRWIPDDDELLRFLASTGTTK
AKLQASPTNLQALWEAHAALAQYFSSATGKPFPHSADQALANVLAFWTGRDYGRIR
LFGRSALGRREKTRAAVNSVNAAVRSQTVGVGQERRSRQVNSRVYSGDIG
ADGHTVMVTEMLQDLWFIHSHSGSALADSETLAVLPATAPORSVSAYVAAAST
IDPKTGDPKXYLPVFNQHEOSPSRSQVDVITWPKNGAVICEAPEQSGQFNTWRFQ
PYKPDQSGSLVEVWEHLAYLPIEAERVRFQNLAHIVQRPSEELPQSXYLITBET
GIGRNLSSVLVRLRGYVAAQIDIQEMLDGTFNGKLSKLLAVVDEAKAGMQYGRW
SHSEKMTNPMERINNEKLEFVEMCCRWLFFSNNDALPFDASDRFRNVANP
THQSPDYPRYIQIANPDPIFAVRKLELLELDISGFKPGAI PMNEAKTKALGAWAS
DLNLLAAPDEWPGPLAERSDVINTLHVSIAKLPGKTIKLEIQAGMKLVERISIG
PNKYRVVVDYTAQOITNPQAYAEAVNARGKYGRQ"
complement(10162..10392)
/gene="mlr0486"
complement(10162..10392)
/gene="mlr0486"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48064.1"
/db_xref="GI:14021452"
/translation="MECVANAGNLVRVMNQSAIVLSIPOVCATLGISRTLDRIYKD
NPSFPRKKIGPLVGLFSLDIETIGSLPDA"
complement(10595..11803)
/gene="mlr0487"
complement(10595..11803)
/gene="mlr0487"
/codon_start=1
/transl_table=11
/product="probable prophage integrase"
/protein_id="BAB48065.1"
/db_xref="GI:14021453"

Query Match

34.6%; Score 114.2; DB 1; Length 346897;

Best Local Similarity

60.5%; Pred. No. 3.5e-13;

Matches 188; Conservative

0; Mismatches 123; Indels 0; Gaps 0;

QY

20 TGATCTTCCGCGACCGCGGCTTCGAGCATTTGTCGAAACCGAGCGGAGTATCGG 79

DB

275086 TGACCTTTCATCGCCCATGACGCGACACATTCGACGTGGATGCCGAAACGCGTCGACGG 275027

/note="ORF_ID:bl15737
putative acyl-CoA dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="BACS1002.1"
/db_xref="GI:27354014"
/translation="NRRAIKGSLALSRVDVLDQAPSTAGEVARIAEQLAPLAAGID
DGSYVPAEVLRFVARGAWGSGHEGPADLRCAIQAAWALGEVCGATAFNAWQNTL
WYAAWNTNKLARGADFGATFGRVLGGTGLSNPMKSPFGIEKLKLGKRKVGEGYIVR
GLAPWYSNIGLPHRYGTIFQCEDDDGIAEGFEGTVMFLADSDPAITLTLPKPFLLM
GTGTYGQRVDVFPDELLIADAPGFPVKTKRAGFILQAGMALGVLDRGCTINIMDEYGV
SPUGHNIRYPOQPVHFRDLAAELAEATMALAROPYNBEETFWRKVIALRLRAGEASV
AAHAAMLHGARGYGLKSHRORRLREAFVVAITYPATKQLRKMLADS"
complement(1191..1733)
/gene="bl15738"
complement(1191..1733)
/gene="bl15738"
/note="ORF_ID:bl15738
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BACS1003.1"
/db_xref="GI:27354015"
/translation="MTAVQKTVLGTGLAPIDKNGLEQLIASGRANPKVVKTLCKRTV
AEGRFHAANYENRLOPYIVDEPGLGGDDTAPNPSLAAAGSLAGVLAHLNAAVHRG
VIWNKLELREGLDGLTAVMGTDGVSDDKPVGFTDVRVYVDNVECEGISOQDEINALVAHV
KWSVPVANTETRFVNLVEGI"
1981..2628
/gene="bl15739"
1981..2628
/gene="bl15739"
/note="ORF_ID:bl15739
TetR family"
/codon_start=1
/transl_table=11
/product="transcriptional regulatory protein"
/protein_id="BACS1004.1"
/db_xref="GI:27354016"
/translation="WAKTSKKISPHAGAGEDEARGLLLSAATHLFCCKGINATGID
AITEAGATKTVLLFGSKTNLVNALESEGWRFVIGAMEQGGGDAQAKLTRIF
PALKSWFAEPYVPGCFINAVEGHDKDAKQFNITALKHKTVLGHIEKLAGELGSEFP
AVLHQGLIIGDAIVAMISRDPGVADTAAITAGTPLLQGSRTKKRAADQLEAV"
complement(2735..4276)
/gene="bl15740"
complement(2735..4276)
/gene="bl15740"
/note="ORF_ID:bl15740
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BACS1005.1"
/db_xref="GI:27354017"
/translation="MTAFSLGTGTFPEVHADRPDKLRKXGLALLVYVEAKGAVAR
DVVATLWPSPQETGRRLRLHRIEOTLGPQFETDTRSAVSPALAKVDTHTLF
ESACRDEPQACRIYRSPDLAGEALDDCPEDDPAFPFRRLARGHLMALRLERLVODK
NACDGEPAATVHAGRVLDLPSEVYGHILRLSLLAGDRSAARHHAALTQRLRDEL
GVAPAEATLMSFPATLAVLDFIVRVKAGVHLVLLAYQAGRAARHHAALTQRLRDEL
WHEFASRTPLASLMKGLRLIVFDRRGIGLSDRVGSAGPIDVTAEIDGTVLRTANRR
AVHPASCEPACIKFANDEPRLVAGLLIFGALAKGWSQDYPHALRASQYDAVRQOQ
VAOQGVPGVITETAPSLADAQARAAMWAGLLPAASPGCISAVLEAFRDADVRHLLAQ
IAPTVLVHRRGDRAVRIAAGRDIASRINGAEFEVLNDGHWFFAGDQQPVLAAGIGGF
MKRTL"
4416..5195
/gene="bl15741"
4416..5195
/gene="bl15741"
/note="ORF_ID:bl15741
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BACS1006.1"
/db_xref="GI:27354018"
/translation="MPARALLXLYAIVSYAAFTVSLYALGVNGVYVVKPSIDVGSPT

ESQH"	543..>1351	/product="16S ribosomal RNA 5'end"
rRNA	Query Match	32.4%; Score 107; DB 1; Length 1351;
ORIGIN	Best Local Similarity	58.7%; Pred. No. 1.6e-11;
	Matches 185; Conservative	0; Mismatches 130; Indels 0; Gaps 0;
QY	10	AAGTCCCGGTGATCTTCGGCGAGCGCGGCTTCGAGCATCTGTCGAAACCGAAGCG 69
Db	101	ATGGCAAGATCATCTACATCAGCAGCAGCGCGCGGCAAGTATCATCGACGTCAGAGCCG 160
QY	70	GGAGTATCGCTCATGGAAGCGCGCTTCTGAACGGCGTGGACGCTATCGAAGCGGTTTGC 129
Db	161	GGCTTCACGGTCATGAGGCGCGCTGAAGAACATGTCGGGCAATCGAGCGCGACTGT 220
QY	130	GGGGCGGCTGTGCTCGGCACGTCGCACGCTTTACGTTGGCCCCGAGTGGCTAGATGCG 189
Db	221	GGCGGCGCTGTGCTCGGCACGTCGCACGCTTATGTGGAAGCAAGCTGGCTCGACAAG 280
QY	190	CTGAACCGCGGAGTGAGACCGCAAGACGAAATGCTGATTCGTTAGCGGAACGTCGGCCG 249
Db	281	ACCGCGGACAGTCGCGCATGGAAGATGCTGATTCGCTGCGCGAGACGTCGACCC 340
QY	250	CATTCGCGGCTGTCTGCGACATCGCGCTTACCGACCTGTCGAGCGCGCTGACCTGGAA 309
Db	341	AACAGCGCGCTGTCTGCGACATCAAGTCAAGCGCGCTGCGCGCGCTGGTGGTGGT 400
QY	310	CTGCGGAGGACAG 324
Db	401	CTGCGGAGGACAG 415
RESULT 8	AE006011/c	
LOCUS	AE006011	10158 bp DNA linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus CB15 section 337 of 359 of the complete genome.	
ACCESSION	AE006011	AE005673
VERSION	AE006011.1	GI:13425254
KEYWORDS	Caulobacter crescentus CB15	
SOURCE	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales; Caulobacteraceae; Caulobacter.	
ORGANISM	Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and Fraser, C.M.	
REFERENCE	Complete genome sequence of Caulobacter crescentus	
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)	
JOURNAL	21173698	
MEDLINE	11259647	
PUBLISHED	2 bases 1 to 10158	
REFERENCE	Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Laub, M.T., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Shapiro, L., Venter, J.C. and Fraser, C.M.	
AUTHORS	Direct Submission	
TITLE	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712	
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA	
FEATURES	1 10158	
SOURCE	Location/Qualifiers	

WDTYDVDAVGMVLDAALYIKNTIDPTLAFRRSCRGVCGSCSNIGGRNTLACTH
GNAVEPKAVQISPLHPNVDLIPDLFLFYAQYASIEPMLHTTTPPQEWKQSPB
DEKLUHTEICACCSSTSCSYHNGDKYLGPAALLHAYRWLIDSDRDEATGDRDLA
LEDFFLYRCHITIMNACVCPKGLNPAKAIKIKOMVERVV
complement (5433. .7223)
/gene="CC3527"
complement (5433. .7223)
/gene="CC3527"
/notes="identified by match to protein family HM"

/codon_start=1
/transl_table=11
/product="succinate dehydrogenase, flavoprotein subunit"
/protein_id="AAK25491.1"
/db_xref="GI:13425259"
/translations="MMSAYKFDHDKFDVVVVGAGSGSLRAALGAAQGLKTAICITKVF
PTRSHVAAQGISASLGNMGDDWMMFDTVKGSDWLDQDAIEYLTRNAPAAVYE
LHWGVFPTDGKIQVARGMTNKGEGPTQRTCAADRTGHMLHTWYGSLAH
DTEFFLEYALDLINDGVCRTAWKLDGDTLHFOAQVILATGGVGRVFSATSA
HTCTGDNVALRAGLPLQMEFVQPHPTGIYAGCLITEGARGGGVLTNBSGFRFM
ERYAPTKDLAPDMVSRAMTEIREGRGVGNPKHIFLHLDLPKILAEPLPGEAA
TAKVFAVDVTKAPIPLVTYHNMGIPTNHYGVVTKSGNPDQVPIPLMAVGEAA
CVSHGANRLGNSLIDLVPFGRAALRCAEILKPGAKQPELKDQVTDALHLPDRFR
NASGSTTAEIRLEKMOQEDAAVPTGESLDSGVARLOAVWVKADLKVSDRGLVW
NTDLMTEFLDFNLIGAVTVNGAANETESRGAHAREDFSDRDAENMKHTLAWLLD
TGKVKDIDRPHSYTMSDIIAYIPPKQRY"
complement (7220. .7639)
/gene="CC3528"
complement (7220. .7639)
/gene="CC3528"
/notes="similar to GP:3273340; identified by sequence
similarity; putative"

/codon_start=1
/transl_table=11
/product="succinate dehydrogenase, hydrophobic membrane
anchor protein, putative"
/protein_id="AAK25490.1"
/db_xref="GI:13425260"
/translations="MSKIDLPQOQRTPLSRAGMGASHGVSHFIFERVSLALAP
LTYGAYALKQVRAHDVTVQMAQPNVAVLLSLILALVHLSAVQVIEDYIER
FTYKALVIGNLFVCLVAGAVFSILKVALTVGAH"
complement (7642. .8049)
/gene="CC3529"
complement (7642. .8049)
/gene="CC3529"
/notes="identified by match to PFAM protein family HM
PF01127"

/codon_start=1
/transl_table=11
/product="succinate dehydrogenase, cytochrome b556
subunit"
/protein_id="AAK25491.1"
/db_xref="GI:13425261"
/translations="MTDTSGLPBPMSPHLOVWMTMACSILHRCGVGLVWGAV
ILAGAAALAAQPDVAYSYVGLLSPGKLVLLGETPAVFNVAITRTQTFWDAGKGF
APRTADTGAFAFAFVAVTIVTWIAGGIGAF"
complement (8221. .8808)
/gene="CC3530"
complement (8221. .8808)
/gene="CC3530"
/notes="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK25492.1"
/db_xref="GI:13425262"
/translations="MSVKDRVRVRETRLLSDNWVLRITTFDWKRDGTWQTSREHY
DRNGAVLLVYNIENVTLLVAVQFPYAFVNGCDLLIEAAGLLDDAEFVRIAEV
EELGVLGRVRFKFAFSPGVTIELHFFVAYEAPAMKRSIDGGHPDEGEDIKLE
MTMEQALMIADGRIDAKTIMLIQHLATVLRTA"
complement (8618. .9936)
/gene="CC3531"
complement (8618. .9936)
/gene="CC3531"

gene
CDS

/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK25493.1"
/db_xref="GI:13425263"
/translations="MPTSLTAYRRLAQRIRPDVAQAARVAEALSRLEADLDAAGEP
GFSEFGKPKSQGVLMGVPVGRKSMVMDLFFDSAPVAKKRIHFHFAWAEVHADID
VWRKGDAAATKARFGOSKGDPAVPTADRIAGEARLLCPDELOVTDIADAMILGRLE
ALFARGTILVATSNRPEDLYKGLARQLFELPIDMLKSALDVAVRGPVDFLRLR
AARTWLAFNKAQAAPFDGJWTDMLDAPETGATLEVLGKMKRLPRAGGLVRSFAS
LCOQALGPQDYLAIARFHTLFLEDPCLTPARRDAARRENTLVDALYEADVKLVALA
EAEPEQIYPREGAFEFERTVSLQENRSADYVVRVD"
complement (9948. .10073)
/gene="CC3532"
complement (9948. .10073)
/gene="CC3532"

Query Match 32.4%; Score 107; DB 1; Length 10158;
Best Local Similarity 58.7%; Pred. No. 1.3e-11;
Matches 185; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 10 AAGGTCCGCGTATCTCCGCGCAGCGCGCTTCGAGCATCTGTCGAAACCGAAGCG 69
Db 3232 ATGCCAAGATCACTATCATCCAGCAGCGGCCCGCAAGTCAATCGAGTCAAGCCG 3173
QY 70 GGAGTATGCTCATGGAAGCGCGCTTCTGAACGGCGTGGACGGTATCGAAGCGTTCG 129
Db 3172 GGCCTCACTGATGAGGCGCGCTGAAGAAATGTCGCGGCATCGACGCGCTGT 3113
QY 130 GGGGCGCGCTGCTGCGCCAGCGTGCACGCTTACCTTGGCCCGGAGTGGCTAGATGCG 189
Db 3112 GCGCGCGCTGCGCTGCGGACGCTGCACGCTATGTGACGAGGCTGCTGCGCAG 3053
QY 190 CTGAACCGCGAGTGAACCGAAGAAATGCTCGATTGCTGAGCGAAGTGGCGCG 249
Db 3052 ACCGCGCAAGTCCGCCATGGAAGAGTTCGATCTCGCGGAGAGCTCGAACC 2993
QY 250 CATTCGCGCTGCTCCCGGATCCGCTTACGACCTGCTCGACGCGCTGACCTCGAA 309
Db 2992 AACAGCGCGCTGCTCCGAGATCAAGTCAAGCGCGCTGGACGCGCTGCTGCGT 2993
QY 310 CTGCGGAGGCGACAG 324
Db 2932 CTGCGGAGGCGCAG 2918

RESULT 9
LOCUS RCFDXE 4151 bp DNA linear BCT 12-MAY-1997
DEFINITION R.capsulatus fdxS gene.
ACCESSION Y11304
VERSION Y11304.1 GI:2094849
KEYWORDS fdxS gene; ferredoxin; ORF E0; ORF E1.
SOURCE Rhodobacter capsulatus
ORGANISM Rhodobacter capsulatus
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Rhodobacter.
REFERENCE
1
AUTHORS Armengaud, J., Meyer, C. and Jouanneau, Y.
TITLE A [2Fe-2S] ferredoxin (FdVI) is essential for growth of the
photosynthetic bacterium Rhodobacter capsulatus
JOURNAL J. Bacteriol. 179 (10), 3304-3309 (1997)
MEDLINE 97294472
PUBMED 9150228
REFERENCE 2 (bases 1 to 4151)
AUTHORS Armengaud, J.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1997) J. Armengaud, Laboratoire de Biochimie
Microbienne, CEA Grenoble, URA 1130 CNRS, CENG-DBMS-BM, F-38054
Grenoble Cedex 9, FRANCE
LOCATION/Qualifiers
1. .4151

gene
CDS

[illegible]

```
gene
/feature="dbdCa"
/codon_start=1
/transl_table=11
/product="alpha subunit of terminal oxygenase"
/protein_id="BAC98956.1"
/db_xref="GI:37651310"
/translation="MTLAPHTPSGRKYNANEIRALVNTDITGRYSPEIYTSBDLYOLEL
BIFPARVCMGHETQIKAGDYOAAYIGEDPVVVVROKDNSIRVFLNOCRRHGRWIC
RVDGNAKATCPYHGMADVTAGNLVSVMEEQEAEGGKLNKAEMGPKOAEVETVKGIL
FANWESAPNDITYLDGAKFYNDIMLDRAGETGVTGQKXVIFPANKWFAAEFGSD
AYHAGTITGSGIAGVAGTIDLSKVQPTDGMNVIIGNHGCGLFSKNPFTYVICG
POLTYLTSVAYEAEVRLGGADVNLCNMVFPNLSFLTGLNTVVMQPRGNEM
ELWTFITLDVADPDNIKEEMRNMYRTFSAAGFEQDDGDNNDIOHVLRGHVARQOP
FNIEMAQHTVQGSTIAPGPPGSPAYTVAESSTGLOALG"
3401..3964
/gene="dbdCb"
/codon_start=1
/transl_table=11
/product="beta subunit of terminal oxygenase"
/protein_id="BAC98957.1"
/db_xref="GI:37651311"
/translation="MLDTPASQAFQVKPELISVELQHEIQFYVWEAKLMADREWEH
FSLMKDKIYNAPLSTSHAREAKLEYTDKGFHPADNMSSLDGRIRKITSYVGS
ENPABRLRHIVGNVLWIDRQADYVDVSTSLNTYRNROEQITDFFAERQDIRRVDD
ACGFEIRKILIDQSTLLSNNISFF"
3979..4311
/gene="dbdCc"
/codon_start=1
/transl_table=11
/product="ferredoxin"
/protein_id="BAC98958.1"
/db_xref="GI:37651312"
/translation="MFNMYLVPLADLPEDGMKVFNGAEPILICNVQGVYAVQDTC
HTMSLCDYLDGHIVECSLHNAKFDVRTGVEKALPACKALKIPKIKENGIEIYDPE
RQTSBGGT"
4312..5139
/gene="dbdD"
/codon_start=1
/transl_table=11
/product="dihydrodiol dehydrogenase"
/protein_id="BAC98959.1"
/db_xref="GI:37651313"
/translation="MCKNPFVALVTGGSGLGKAIIVERFLAEGAKVAILDRSEERVOEV
VAMFGGVVGTSGDVRSMADNKAARVCEAFKIDITLVNAGVWDYNASIAGTSEEK
LABAIDELVGINLKVILAAKALPELYKGNALFTVSNAGFPYGGGVLYTASKHG
VVMIKOLAHAWAPHIRVNGVAPGGIGTISRGFPSSLOEGTFRFSLPDLDDMKWPL
EKAFVASEVTGCVFFANRKNQDPATGSLVNFDDGIGMRFMSPNWKGELDEKFGA"
5233..6407
/gene="dbdE"
/codon_start=1
/transl_table=11
/product="salicylaldehyde dehydrogenase"
/protein_id="BAC98960.1"
/db_xref="GI:37651314"
/translation="MSELLRVGDEGSAADCAVLERSPADGVVVSVRAAKRADVE
AAIATAAYIEMSKTSPSHRMLLKADELEKRAQSLVEMKDEVGSAELWAFNV
MAAEFLPRAASMATOLOQETITPTDKPGLTSLRQPVGVLSIVPMNGPVLAAARI
AYPLVCGNTVIFRASMPRVLAHVAAGFPQDSIKLVNLPEDAPALIEAMIA
DPRVSVNFTSGTRGRIIGELCGRHLKXSVLELGDKSPVLVLCADYEAANVAIAG
AEFYQOICMSTERLIVHETIADEFVERLAARVMTLKYGDPRQAQAAWALGPVVSIGAA
DLNLRLRLRLDDAVSKAKIAGGAAQQTENDATVLDHVTDRDAIYREEVFGPIL"

gene
32.0%; Score 105.6; DB 1; Length 6407;
Best Local Similarity 58.7%; Pred. No. 2.6e-11;
```

ORIGIN

Query Match

Best Local Similarity 58.7%; Pred. No. 2.6e-11;

```
Matches 183; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
Oy 19 GTGATCTTCGCGCAGCGCGCGCTTCGAGCATCTGCTGAACACCGAGGAGTATCG 78
Db 429 GTAAACATCATTAACATGAAATGACGCAATGACGCGACCGTGGAGCGCATCCGGGTGCTCC 488
Oy 79 CTGATGAAGCGCGCGCTTCGAAACGCGTGGAGCGGTATCGAAGCCCGTTTTCGGGGCGCGC 138
Db 489 ATCATGGAGATCGGGTGAACACGCGGTTCGGGGTATTGATGCCGAATGTGGCGCGCGC 548
Oy 139 TGTGCTTCGCGCAGTGGCGACGTTTACGTTTGGCCCGCGAGTGGCTAGATCGCTGAACCG 198
Db 549 TGGCGCTTCGCGCAGTGGCGACGTTTACGTTTGGCCCGCGAGTGGCTTAAAGAACTTCCCGCG 608
Oy 199 CCAGTGTAGACCGAAGACGAAATGCTCGATTGGTATGCGGAGCAACGTCGCGCGCATTCGCGG 258
Db 609 CCGTCCGATTCCGAGAGCGGATGCTGGGATTCGCTGCGCATCGCAGACGACACGCGC 668
Oy 259 CTGCTCTGCGCAGATCCGCTTACCGACCTGCTGCGAGCGGCTGACCTGGAACCTGCCGAAG 318
Db 669 CTGCTCTGCGCAGATCCGCTTACCGACCTGCGAGCGGATCGTCTCAACGCGCGAA 728
Oy 319 GCACAGTCATGA 330
Db 729 TTCAGTCTTGA 740

RESULT 11
BX572605/c 349746 bp DNA linear BCT 18-DEC-2003
LOCUS Rhodopseudomonas palustris CGA009 complete genome; segment 13/16.
DEFINITION BX572605 BX571963
ACCESSION BX572605.1 GI:39650627
VERSION complete genomes.
KEYWORDS Rhodopseudomonas palustris CGA009
ORGANISM Rhodopseudomonas palustris CGA009
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiacae; Rhodopseudomonas.
1 (bases 1 to 349746)
Larimer,F.W., Chain,P., Hauser,L., Lamerdin,J., Malfatti,S., Do,L.,
Land,M.L., Pelletier,D.A., Beatty,T.J., Lang,A.S., Tabita,F.R.,
Gibson,J.L., Hanson,T.E., Torres y Torres,J., Pares,C.,
Harrison,F.H., Gibson,J. and Harwood,C.S.
Complete genome sequence of the metabolically versatile
photosynthetic bacterium Rhodopseudomonas palustris
Unpublished
2 (bases 1 to 349746)
Larimer,F.W. and Harwood,C.S.
Rhodopseudomonas genome consortium
Direct Submission
Submitted (24-JUL-2003) Submitted on behalf of the Rhodopseudomonas
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov
Location/Qualifiers
1. 349746
/organism="Rhodopseudomonas palustris CGA009"
/mol_type="genomic DNA"
/strain="CGA009"
/db_xref="taxon:258594"
96..521
/locus_tag="RPA3709"
96..521
/locus_tag="RPA3709"
/functions="InterPro IPR000971 COGs COG1017"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible hemoprotein"
/protein_id="CAE29150.1"
/db_xref="GI:39650628"
```

/translation="MMTPQAIALVQOSFARLTPISTDETATLTYERLFPATAPELRPLFH
 GDMRQCKLMLSTLVVVTGLTDLTTLTPATGLRLARLHVAGVSAHYTPFGAALLWT
 LERLGAETWPELAAARDYAMLSETMLAAAYGDTTPS"
 564..2336
 /gene="nirA"
 /locus_tag="RPA3710"
 564..2336
 /gene="nirA"
 /locus_tag="RPA3710"
 /EC_number="1.7.7.1"
 /function="InterPro
 IPR001424:IPR005117:IPR006066:IPR006067 COGs COG0155"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="possible ferredoxin-nitrite reductase"
 /protein_id="CAE29151.1"
 /db_xref="GI:39650629"
 /translation="MGSEFSIDQKRYLEGFATVTAARAAGSLPAGSAASPGSDAI
 HQQADRTASGKLNQKRWQHPFDAYARLQQAQKDAAPKPADNFRWRYGLF
 YVAPQTSYMCRLIPNGILTSWQAGLAYLADTLADYAHVTTTRANLQWREIAPDA
 VQLQGGIESLGARSGADNIRNVGSAIDPQELIDTRAYARDWHFYILNERAL
 FGLPKFNVAFDGGGLIPTLEDNTDIAFOAVTIGDGGVAPGVYFRALGIGTHKDF
 ARDTGVAVPADADAVURVPIEHGDRDTRNKAKELKYVLDAGFDPKYLEVEDKL
 GRKLIRVPAEAAAPRQDRAHLGVHPQQQAGNWNIGVRLATGHLTSVVMRGLAIEIA
 RSPGQDRLTIVWONLLISGVPDARVDATAALGALGALTASPLRAGLACTGASGC
 RFAAHTKETAEIAQYCEPRVPLDTPINHLTGCHNSCACHFISDGLIGAKVAVSD
 EDSVEGTHVHGGFGVDAAIGTEVLDRDKQDDAPRIERMLOTYLARRASPAETFLA
 FVRQDAALQQLPAYEQSPETVA"
 2333..3952
 /locus_tag="RPA3711"
 2333..3952
 /locus_tag="RPA3711"
 /EC_number="1.8.1.2"
 /function="InterPro IPR001433:IPR001709:IPR003097 COGs
 COG0369"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="possible sulfite reductase (NADPH)"
 /protein_id="CAE29152.1"
 /db_xref="GI:39650630"
 /translation="MSQNLOPPIPLLVPEYAPFSEQRWLNGFFAGVLSDQGGVTA
 LSQQQACDRNYSAGPAIDQGEPRDDGGAPWHQDSLPTEGMRQLABGKPLRWKMAAQOD
 CPGQVYALVAGGAIHQEPRNLNLCAPGGKDTARMVKTLTETQIGRAPTIONAPSIAA
 DAAPVVALPRTSRDNPATAKVLRSKLNKPGSEKETHVVELEDCLSYEAGDSFG
 LFTNDPALVDVHALGAPAEFFIAQSLNQLTLDVSVSPAPDMFLQISITIGSD
 RKKARSASGEDPDGAATLDVLALEKFPGLRPDPPEAFIEALDPLQPLRLYISSPN
 KTTPCRLSLTVCVRYAIGKQRLGVCSTGLAERVRPETIRAYVQKAHFLPSIDPN
 QPIIMIGPTGIAPFLAFLHQRQIAQPKNWLFFHQSGASDPFFYEDELKAMKTAGH
 LTRLTAWSRDSGEKIVQDRIREVGRDLMSWLAEGASLYVCGDAKEMAKDVERALVD
 IVAOHGARSAAEATAFVSELKKQGRYQQDVTY"
 complement(4092..4863)
 /gene="pimE pseudogene"
 /locus_tag="RPA3712"
 /pseudo
 complement(join(4092..4487,4486..4863))
 /gene="pimE pseudogene"
 /locus_tag="RPA3712"
 /EC_number="1.1.1.1"
 /function="InterPro IPR002198 COGs COG1028"
 /note="frameshift"
 /pseudo
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="short chain dehydrogenase"
 /db_xref="PSEUDO:CAE29153.1"
 complement(5013..6155)
 /gene="pimD"
 /locus_tag="RPA3713"
 complement(5013..6155)
 /gene="pimD"

/locus_tag="RPA3713"
 /EC_number="1.3.1.62"
 /function="InterPro IPR001687:IPR006090:IPR006091 COGs
 COG1960"
 /note="observed by proteomics
 Citation: Proteomics from VerBerkmoes et al. (2003)
 unpublished"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="pimeloyl-CoA dehydrogenase (small subunit)"
 /protein_id="CAE29154.1"
 /db_xref="GI:39650631"
 /translation="MDFDLSEQRLLKESVEGLLKGSYDFDSRKYAKEKGSRAVW
 GKFAEQGLIGLPFSEEDGGFGAGAVETMVMVEALGHSVLBPYLPVTVVIGGFLRRAG
 SAAQKAHLPGIIDGSKTFAQALEKNSRWLDGVSTAKSGDGMVIDGEKFLVNLG
 EAADTLVYARTKGGORDTGTVGVFLVPADAKGITRKGYPTQDGLHAADIITFGVQVG
 ADAAIGDPENALLEIEAVDDARTALCAEAVGLMDESITTVYEIKTRKQFGVPIGSF
 QYLQRAADMVFAEQASMAFMATMAAEFDDAKERAGATAAKKVOIGSKGFVQQQS
 IQLHGIGMTWEAKIHYFKLTMTEIQTFGDTDHLARVSAGGGLI"
 complement(6324..7514)
 /gene="pimC"
 /locus_tag="RPA3714"
 complement(6324..7514)
 /gene="pimC"
 /locus_tag="RPA3714"
 /EC_number="1.3.1.62"
 /function="InterPro IPR006090:IPR006091:IPR006092 COGs
 COG1960"
 /note="observed by proteomics
 Citation: Proteomics from VerBerkmoes et al. (2003)
 unpublished"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="pimeloyl-CoA dehydrogenase (large subunit)"
 /protein_id="CAE29155.1"
 /db_xref="GI:39650632"
 /translation="MDLNFSEKIEAFRDEVRQFQKDNVPKTRQKLEGRHNTKEEMV
 EWYRLNKGMVATHTWPKYGGTGSVQCHIFNEELQAAPQAPLQAGVSNVGVYI
 TFGSEQRKRLPLRIANVDDMWQCGFSEPGSGSDLASLTKAERKGDWINGQKWT
 TLAADHAIICLRTDPAKQOEGISFILVMKTKGITVTRQTTIDGHEVNEVFDD
 VVPLNENLVGKNGDYAKFLQGNERTGIARVGNKERRIRIKQLAAQVSEGGKPI
 EDPKRDILAAVEIELKALELTQLRVADEQHGKGNPKNPASSVLKINGSETQOATTE
 LMVEYIGPFAAPYDVHGDDSDNTWDTAQIAPGVFNKRKVSIVYGSNEIQNRICKA
 VLGL"
 complement(7729..8916)
 /gene="pimB"
 /locus_tag="RPA3715"
 complement(7729..8916)
 /gene="pimB"
 /locus_tag="RPA3715"
 /EC_number="2.3.1.9"
 /function="InterPro IPR002155 COGs COG0183"
 /note="observed by proteomics
 Citation: Proteomics from VerBerkmoes et al. (2003)
 unpublished"
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="acetyl-CoA acetyltransferase"
 /protein_id="CAE29156.1"
 /db_xref="GI:39650633"
 /translation="MTEAVIVSTARTPIKAYRGALNATEGATLLGHAIEHAKRAGI
 SVLPFDGVVLAAGGGESISLVNDKMTFHAVDPALEAIKGDVYVAMLDTAETAKRY
 DPKEVEDVMGAAPQCGATGGNIARAKALLRAGLPVTAGTITDRCQASGLQALAAAR
 SVLPFDGVVLAAGGGESISLVNDKMTFHAVDPALEAIKGDVYVAMLDTAETAKRY
 GTSRRQDEYSLSESORRTAAQGGKFNDEIAPISTKMGVVDKATGAVSFKDITLSQD
 EGPRTETABGLAGLKAIVRGEFTTAGNASQLSDGASATVMSKTAAGKLEPLGI
 FRGMVSYGCEPDEMGIPVFAVFRLLKRLKHLGSLVDIGLWELNEAFVQVLYVCRDKLGI
 DPEKLVNNGAISVGHPIYGMGARLAGHALIEGRRRKAKYAVVTTCVGGMGSGAGLFE
 IVH"
 complement(8948..10606)
 /gene="pimA"

```

CDS
/locus_tag="RPA3716"
/complement(8948..10606)
/gene="pima"
/locus_tag="RPA3716"
/EC_number="6.2.1.-"
/function="interPro IPR000873 COGs COG0318"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="pimeloyl-CoA ligase"

Query Match
Best Local Similarity 30.5%; Score 100.6; DB 1; Length 349746;
Matches 157; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 55 GTCGAACCGGAGGAGTATCGCTCATGGAAGCGCGCGTCTTGAAACGCGTGGACGCT 114
Db 280890 GTCGAAGTCGAGGAAGCGCGACGGTATGGAAGCGCGATCGCAACGCCATCCCGCGC 280831

QY 115 ATCGAAGCCGTTTGGCGGCGCGCTGCTGCTGCGGCACGTCGCCACGTTTACGTTGCGCCC 174
Db 280830 GTTGAAGCAGAGTGGCGCGGCTGCTGCGCTGCGGACCTGCGCACGCTCTATGCGACGAA 280771

QY 175 GAGTGGCTAGATCGCTGAAACCGCGCGAGTGAGACCGAAGACGAAATGCTCGATTGCGTA 234
Db 280770 GCCTGGCGGAGAGGTGCGCGCGCGCTGCGCGATGGAAGAGACATGCTCGACTCGCG 280711

QY 235 GCGAAGCTGCGCGGATTCGCGGCTGCTGCTGCGAGATCGCGCTTACCGACCTGCTCGAC 294
Db 280710 TAGACGTGCGCGCGAAGTCTGCGGCTGCTGCGGATCAAGTCTCGAAGCAACTCGAC 280651

QY 295 GGCTGACCCCT 305
Db 280650 GGGCTGATGCT 280640

RESULT 12
AP005215/c
LOCUS AP005215 309350 bp DNA linear BCT 10-JUL-2003
DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 2/11.
ACCESSION AP005215
VERSION AP005215.1 GI:23492123
KEYWORDS Corynebacterium efficiens YS-314
SOURCE Corynebacterium efficiens YS-314
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
1 Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikee, K. and Gojobori, T.
Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of Corynebacterium efficiens
Genome Res. 13 (7), 1572-1579 (2003)
22723752
12840036
2 (bases 1 to 309350)
Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.
Direct Submission
Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan
(e-mail:bioinfo.go.jp; Tel:81-3-3481-1933, Fax:81-3-3481-8424)
Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan
Nakamura, Y., Ikee, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Itoch, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan
```

```

FEATURES
source
1. 309350
/organism="Corynebacterium efficiens YS-314"
/mol_type="genomic DNA"
/strain="YS-314"
/db_xref="taxon:196164"
/complement(163..462)
/notes="CE0288"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17098.1"
/db_xref="GI:23492124"
/translation="MMDPNSAALYLRVDPWPCAREARFSLHCGIMTGLLVVLIPLLL
FFLPAMQRIEISLLSTQDVNSRDVTTSSDDASAGAGSGEDGPOSQPKAA"
437..1231
/notes="CE0289, similar to AL583925-21|CAC31817.1| percent
identity: 66 in 240 aa"
/codon_start=1
/transl_table=11
/product="putative endonuclease III"
/protein_id="BAC17099.1"
/db_xref="GI:23492125"
/translation="MAAEFGSITPLTRPRYGVSHIANGGETDGRKERRINRTLAV
AYPAHCELDFTNPLTVAITLISAOCTDVRVQVTPALFRYPTAWDYANADRAELE
ELIRPTGVTKATSLIGLGRALVSLDGVPHTELBELKLGIGIKRTANVVLGDAFG
VPGITVDFHGRLARBLKLTETDTPVRVEHEIGALTEKEMTLFSLHLIFHGRRICHS
RRAACGACMLAADCPSPGLEPADPMEAQKLKSDREHLIKVAGM"
1237..1815
/notes="CE0290"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17100.1"
/db_xref="GI:23492126"
/translation="MTTSARSLVGVVVVAVLVAVIPQLLSGTTGCGDATGAAGDG
AGEASTIVADRPDCQDGAAGVQLPCGEGCGNDLPTVVNVVAMWCEPCRDPLPV
FDFEAFSPHLYVGHADTNAANGAAMGLDGLVPSYQDDSNLFGATLGLPAVPI
TVVVDPDGELVGTFFRTFDSVQDLEAAVAGL"
1812..2579
/notes="CE0291, similar to AL022121-21|CAAL7994.1| percent
identity: 42 in 217 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17101.1"
/db_xref="GI:23492127"
/translation="MMPRHTPNHLYPDLPHDPGENTELAPAKAPVWRRMLMDRTEA
GHENPLACETVGTTHAEKQAAVLPLFSGSETSPDLNDASVLTTHRSPTWESHGQ
IAPFGRIDEDINAVDCAPRAWEETGDRRTATPLAQLEEVHIRATGYVYPIGQH
WHSPFVAVVSPDEADVFDPAYELIDPANRLMWGRWKKGPAFRINGVYWGFTGG
LLSAIDQAGWATDWDTRIYDLEDTLATSRNNEPLR"
2737..3933
/notes="CE0292, similar to AF322256-34|AAL15613.1| percent
identity: 35 in 396 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17102.1"
/db_xref="GI:23492128"
/translation="MLASIPVDIIIGIILVFWLWSGRCQAFISLIATIGVAVGLVI
GAAPVAMGLTSPDRFLATGTVVLVGLNGLGHAIGHAIKRGIRFRSSRFLDS
LVGSIQVATLIVAMVATPLATMGVTASGIRSEIILAFVDDNTPRGLETLARI
AMLTGFLPLVSPDNQPVREAPNTINVSVALVEOMRPSVVMGDCAGCSREL
MGSGFVAAPDYVITINAHVACTDAVNLDTVVGRAAEVYFDPDAIALVHSPDLGLE
ALPUSDPLNSGADALVMPGSPGDPASPARVRERILITGANIYSGQHEREVSVR
GNTRSGNSGGPMVDTGEGNVVGVVGAAGDSTGFTVLTAEVERRIGLDLFTPVDT
```

Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8881 Japan
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.

LOCAYS"
complement(3955..5025)
/note="CE0293, similar to AX065881-1|CAC26180.1| percent
identity: 56 in 305 aa"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="BAC17103.1"
/db_xref="GI:23492129"
/translation="MTTRTSFGSPGVHACQSPASPSRSMVIRKFTPEPWKLWMTA
LRRRSNVAASVETVAVVAGVPYRHHVYVGRVLRHVAVAGSPADPLVLLHGA
FGGVYKDVIAAGLAAFFVAAVDLGYGLSDKPPSGYDIIRRSAGDINGVILGHD
DAIVGDTGSLAWASTLYLPDRAAGVILSIGAVHPADLRARLRKPHLFGSLARIA
LFSPIVPLASIRVPRASREVTRCTTASVORSNAYTCATRLKALALDHTFPI
VTRNLMTASLPGSDHOLACPCVLLTPRDNTEYLASLARSSTGPFPLVSIPIIT
ALSTLEPLAFTEMIAFLSDS"
complement(5229..5895)
/note="CE0294, similar to AE001940-6|AAF10446.1| percent
identity: 23 in 150 aa"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="BAC17104.1"
/db_xref="GI:23492130"
/translation="MARLGLKNAQVIRLKEHWSFQVNNNDGLFTDGTTFSPKVNAI
PLSDVTSISGEASIGITLISNATQMSLSLFAEVELAGEAKKAIGGAGFAA
GVIALYSFPFPLATLKIWLDSWLSYLVLEMLLIAAVALFGKVKVKGWAPK
KTIQSVNLXNLVPGQATSNLERSDKRLYTSNFAHPSGTAASSTIIPSGHSTAVRK
"
complement(6235..7683)
/note="CE0295"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAC17105.1"
/db_xref="GI:23492131"
/translation="MSGSDDLFPETLQRPVAPDPVAPVUKRPAFNPFAVMLLP
PSLDELPHDLHAFIADLVDELGLTGYDSYTRSKRPPYDFRMLRVMLYGYCTG
VRSRKLCTDISVAFRLWLAGADPDFRAFSPFRARHLQALAGVFFVQALALCREAGM
VLCATVADLGTQVANSRRKMSYQRLVPAEKLAEQVQMLNDAATDEVEDQRYG
VDARGDELPELRSARLARLARQOOLEADAKARQVATKARDRGDDDTVDVK
QAAADTAVPVLAQORNPDPDARIKMLSSGPDYCVNAOTVDAEHQVIVATELNT
ADTGATVPMVLAETQOLGWPAWELMDAGYCKKNLTVEAGLEDHTEFFATLST
KHSEKIDAPAGRPIDNATLRMRGRRLRYKRGKAVYKRSVIEPVFGQIATRQGH
VVLKGLAARAEWDLVAGCHNLLKLSFRQTA"
complement(7804..8850)
/note="CE0296, similar to AJ010584-1|CAB55344.1| percent
identity: 31 in 273 aa"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="BAC17106.1"
/db_xref="GI:23492132"
/translation="MLGEVGDGAERGWLHIVNSPPQGGRIETPTSTRNHTSHR
NAGLPHIHRKTRVTCVTPNDFSPGHFNVAQDSTRAAPRRVAFDLDITIAM
SSTVAYGREFMQNGLISPVTAQLSLAQATVMFAGHTSEOMDTRDOLTAMVRGWDVQ
QVRTAETWCTVPTPIYABARELIDYHREGRHDVILIASVKEAVEPTARELVNQ
TVSTLTIDGRYGEVLFYCKGAQQAQVNLADQGYDESSYAYSADATDIPMD
IGVHPVAVNPDRALKKAALRGWDILSFKNPEPLFQMPSTREMGIGTVGVVAGIAAVTA
GGMWARRARGSA"
8903..10303
/note="CE0297"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAC17107.1"
/db_xref="GI:23492133"
/translation="MQQPDILCGLPFRPPEVGLPTARHPPRGAGRPHPSPSSIGPL
PGGLPGGRPPGYPGPGSPKPGVTHRLAASLRAGPAASRVGMGNAATHTTTR
TAGTARSTQSDOAILVAVEDPVLPHEAMHVAATGTAIVDDTPDVIHRHLKYSA
VLIDAPTAQITNDRERDIPLGSDPGPDYHTALAIKRAQALLPAQTAELIQLG
REDSPAPGRHATVTGVLGAGGVGVSTIAALARISAHTRLVLDVAVPTSGGIDL
LVGAETIGARWPDGLGTRGAVKAEDVLALPVMDETMFIUSGARSVPVGTDFDLGPD

VTALTCTAAGGELEVVVDLNPGBITREVIPLDLHLILVVPAAEVRAVAAAERHUL
RAFPVPSVLRHRCWGLDIEVERILGTFTVIAELGTITRPLRAVENHGLTGLPRP
LVTGNIAAAEIRGRG"
10303..12111
/note="CE0298, similar to AL035636-18|CAB38493.1| percent
identity: 55 in 348 aa"
/codon_start=1
/transl_table=1
/product="putative conjugal transfer protein"
/protein_id="BAC17108.1"
/db_xref="GI:23492134"
/translation="MSTTIPDELEQVORVLAGLPEHPTSPVVOVIREQAGVISDE
EVALRLRESVGIGLEAALAPGVTVLNGPRTVWIDRGAGLSQLDLGSEE
AVRLAARLQCGPRLDAAQPFVDDGCTTRDDGSTIRHAIPLADSGTCLSWELLR
QASRLDLVAGGTLTGTRAVLRITVAQRRAFLVVGTSKTKLLAMGEVAPDQ
RICTEDTAELNPPHPTGNLVITRSANVEGAGATMSDLLKQAMRRDRIVLGEIRG
AEVVELLAALNTHEGGAGTTHANSIAEVPARMEALALGLADLHLSQALADVV
IVMRTTTRRLAQLGLVLAIRSRMWCCTGTNTWMTQWRHGFTRDPPFCRGDGA
GAGAGGTGRRREASGSGTPATCGHRRRPRVRGPRVRRHRRPRHGGGAADRRTRGELA
"
Query Match 27.0%; Score 89; DB 1; Length 309350;
Best Local Similarity 60.4%; Pred. No. 3.5e-08; Indels 3; Gaps 1;
Matches 165; Conservative 0; Mismatches 105;
QY 25 TTCGCGCAGCGCGGCTTCGAGCATCTGTCGAAACCGAGCGGAGTAGTCGCTCATG 84
Db 278820 TTCACCGATACCACCGGAGAGCCCGACCATCAACGCAACGCTCCGTGATG 278761
QY 85 GAAGCGCGCTTCTGAACGGCGTGACGGTATCGAAGCGCTTTCGGGGCGGCGCTGTC 144
Db 278760 GAGACGGCGCTGCGCAACGGTGTGCGGGCATCTGCGCGCAATGCGGCGGTTCCTCTGCC 278701
QY 145 TCGCCACGTCGCGCATCTTACCTTGGCCCGAGTGGCTAGATCGCTGAGAACCGCGCGAGT 204
Db 278700 TGGCCACGTCGCGCATCTTCTCGTCCAGCCGCGCGACTTCGAGACGCTCGCGCGAGGAG 278641
QY 205 GAGACGGAAGACGAATGCTCGATTGCGTAGC---GGAACGTGCGCGCATTCGCGGCTG 261
Db 278640 GAGATGAGAGACGAGATGCTCTGGGGCGCGCGCGTGGACCGCGAGACTGTTTCGCGCTG 278581
QY 262 TCTGTCAGATCGCGCTTACCGACCTGCTCGAC 294
Db 278580 TCTGTCAGATCGCGTTCACCGAGGCGATGAC 278548
RESULT 13
LOCUS
DEFINITION
AB008062 10495 bp DNA linear BCT 18-DEC-2001
Agrobacterium tumefaciens str. C58 circular chromosome, section 120
of 254 of the complete sequence.
ACCESSION
AB008062 AB007869
VERSION
AB008062.1 GI:15156410
KEYWORDS
ORGANISM
Agrobacterium tumefaciens str. C58 (Cereon)
Agrobacterium tumefaciens str. C58 (Cereon)
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE
1 (bases 1 to 10495)
Hinkle,G., Slater,S.C. and Goodner,B.
Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
Unpublished
JOURNAL
2 (bases 1 to 10495)
Hinkle,G., Slater,S.C. and Goodner,B.
AUTHORS
Direct Submission
TITLE
Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street,
Cambridge, MA 02139, USA
JOURNAL
Location/Qualifiers
FEATURES
1..10495
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/molecule_type="genomic DNA"
/strain="C58"

```
gene
CDS
/db_xref="taxon:181661"
/261..629
/genes="AGR_C_2493"
261..629
/genes="AGR_C_2493"
/notes="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2493p"
/protein_id="AAK87141.1"
/db_xref="GI:15156411"
/translacion="MAAIVSIFEEAPDNVGGAPAVAKKPIDFDHLAQQTGMGDKOLEIEV
LQIFRSASAAALHEMGADNKTVTATAHLKGSQAQVAGASAVSAAAAAEEKGNDSAL
IAKLAIAIVEAENFIKLGR"
648..1079
/genes="AGR_C_2495"
648..1079
/genes="AGR_C_2495"
/notes="FERREDOXIN, 2FE-2S (FDII)"
/codon_start=1
/transl_table=11
/product="AGR_C_2495p"
/protein_id="AAK87142.1"
/db_xref="GI:15156412"
/translacion="MPFCAGSACVSRGSLPRVGLSMCLQFQSFSDHILDSGNTKLTV
AFDGTHEUDVNGSTVMEVNAVSVPGIDAECCGACACATCTHYVDDAWAERVGGPE
PHEEDMLPAFVRPTSLSCQIKNDDELGLVHVPRQG"
complement(1386..2012)
/genes="AGR_C_2496"
complement(1386..2012)
/genes="AGR_C_2496"
/notes="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2496p"
/protein_id="AAK87143.1"
/db_xref="GI:15156413"
/translacion="MTKSSSALCAVEFAGLALSLSGPAETIVRKYTSYFSVSGKTA
ALDQQLSKHGLTENTGARGHGAIEKFGELTYKNGVCSVGTAKVLNTRLLIP
QWNRRTTAEAFIWTLLADIKRHEENHABIAHRSUERQTSURPORSCEVLQ
GMWGIQTQVNDHRDQRFVDVESKPNPDMRTLLKNRLEQQARR"
2117..2577
/genes="AGR_C_2497"
2117..2577
/genes="AGR_C_2497"
/notes="(U72862) dihydroxyterate synthase"
/codon_start=1
/transl_table=11
/product="AGR_C_2497p"
/protein_id="AAK87144.1"
/db_xref="GI:15156414"
/translacion="MITVGSNAWQAAGHSLKLDGRGRIAMVATPDSPFSDGGRYLA
VDAASHATCYEEGADIIDIGESTRPGAAVTEAEQCDVLPVIEKLCRETDLVLS
VDTYRAATARLAIGAGAHVNDVFLGDKDPNAGVVSSTRAGICIMHTRDRQKLADV
IDQFEFLNLEIAEAGIALDAVVDLPGFGFAKDESENVALMARFGELAAFGLPVL
AGTSRRKFTGSUTGRDVAESERDGTAAITAILRLAGASIFRVHNVAAATRDALATADV
LAKASAKADA"
2931..3422
/genes="AGR_C_2498"
2931..3422
/genes="AGR_C_2498"
/notes="DIHYDROEOPHTERIN ALDOLASE (DHNA)"
/codon_start=1
/transl_table=11
/product="AGR_C_2498p"
/protein_id="AAK87145.1"
/db_xref="GI:15156415"
/translacion="HFMRFQKRRQBPMPRSTSSAGMPCPEDGAARFRKQSCGGDAMS
RTIILKNSFFARHLEQEEVIGORFFVDAEMEVEAGDALETDIDIENTVDYGVAFV
VLEKSVGGRRYLIALANDIAKALRARYPQIWLKITVRKPSAPVGLDYAQVSE
HRA"
3481..3942
/genes="AGR_C_2500"
3481..3942
/genes="AGR_C_2500"
/notes="2-AMINO-4-HYDROXY-6-HYDROXYMETHYLDIHYDROPTERIDINE
PYROPHOSPHOKINASE
(7,8-DIHYDRO-6-HYDROXYMETHYLPTERIN-PYROPHOSPHOKINASE)
(HPPK) (6-HYDROXYMETHYL-7,8-DIHYDROPTERIN
PYROPHOSPHOKINASE) (PPEK)"
/codon_start=1
/transl_table=11
/product="AGR_C_2500p"
/protein_id="AAK87146.1"
/db_xref="GI:15156416"
/translacion="MARALRELDADHDCRQVAVSGLYRTPPMGKTDQADFFNCALVE
TSITALALQLCLDIERGMKRVSTERWGRPTIDIVLTGNEAIVTESIEVPHRMT
RAFLVPLADIAPLEVKGPVQWQQADRSGIVRANEKREWTLPLGDD"
complement(3957..4517)
/genes="AGR_C_2502"
complement(3957..4517)
/genes="AGR_C_2502"
/notes="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2502p"
/protein_id="AAK87147.1"
/db_xref="GI:15156417"
/translacion="MYSRFRFSLMGLAAILSVGGFGMSDVHAASDRDDVFRSVAG
SWKPGEIVAGKYKGTFTCDLTGEPLDKQTKIGLGTCTRVGFQSPMSAVISQKN
SYRGKFLDGAAGKGLDITTCQVSGDKVWVGINRAKLGAMVARVSDNSNMTISVKV
ADQMIPVIGVTLARDIDQIAVGSIKP"
complement(4647..5726)
/genes="AGR_C_2505"
complement(4647..5726)
/genes="AGR_C_2505"
/notes="(AY008264) hypothetical protein YcJf"
/codon_start=1
/transl_table=11
/product="AGR_C_2505p"
/protein_id="AAK87148.1"
/db_xref="GI:15156418"
/translacion="MKAPQNDPQTRRAAFTLEBAAAPSAQTQKRAPSPDAEISL
TPDEDPPFADIDIAALFVATPPKSRFSFGKLGALGVLSLAFGLWADQLIENL
FKSRDMLGYTATIALIYALVAVLQVREVFGLMRLNAVQSLKADAEASLDKSPKA
RAIVTRLNAVLSHRAETAKRAALKTENDVIDGPHLIELARELELVPLDRQARAIL
NSSKRVSVTVAVSPRAVDLAVLVFVTRLVRAVAELYGRPGTGLMKLLRDVVAHL
AVTGSIAVGDLGAQVGLGHGLASKLSARLGEVINGLMRTARIGIAMDLCLRPPLPRAY
KRFGLDPMSDLTPDLSSGKNGKA"
complement(5723..7111)
/genes="AGR_C_2507"
complement(5723..7111)
/genes="AGR_C_2507"
/notes="(AY008264) hypothetical protein YcJx"
/codon_start=1
/transl_table=11
/product="AGR_C_2507p"
/protein_id="AAK87149.1"
/db_xref="GI:15156419"
/translacion="MRLGVTSRAGKTVFISLVHNLNGGRLPLFAMRSORVSVN
RLEPPDDAIPTQYEDHICALVRDLWPDSTRASELRTILDYQASGQWFSAGK
LSIDIVYDPEWLDLPLSODYKQPSDATVALAGSGIRAEIAQEWLAISLDINAP
ADETARRLAESAFAAYLACKSDESLTLPGRFLMPCDLSGSPALTFAPLPGVTD
KAPGSRAMERRIDAYSIVVVFREHFAKDLQIVLIDTLQAVNRGPAAVDLE
RALGDVLACFRPTNSILSLRIRRDVKVLIATKADHLHSHSHDLERLITFDNRT
ITTTGMNGAGIEYVALASRTEASVQDGHLPVIVGTWAGETNGEIPDNKRT
AIFPGDLPEPFLFRSIDQDGKATLPDVNVVVRFPNIDEPGSGGIRLSVPHILD
RAMQFLFGDKLA"
complement(7330..7839)
/genes="AGR_C_2511"
complement(7330..7839)
/genes="AGR_C_2511"
/notes="hypothetical protein SCI28.03c - Streptomyces
coelicolor"
/codon_start=1
/transl_table=11
```

```
/product="AGR_C_2511p"
/db xref="GI:15156420"
/translation="MTASPTRYLLRHAAWAAAGERDFRGLNEAGFAEELIAD
LAADRRPDLISSTAAKCTQAWRAFNEGIDIVIDENYNARSETYLISLIAAQ
TEQSVMLVGNHPTMEATLEAMIGEDLLHAALFSGPFTSLGLAVLDQDDSAASGNRWR
LIDFLAFGK"
8082..8528
/gene="AGR_C_2511"
8082..8528
/gene="AGR_C_2513"
/notes="DNAK SUPPRESSOR PROTEIN HOMOLOG"
/codon_start=1
/product="AGR_C_2513p"
/db xref="GI:15156421"
/translation="MRCSSRIALSEKIDSKVLSDEDFMASORAYPRAXLIAWK
NDILREARETLGLAESANHPDLADRASSETDRAIELRARDQRKLKIDLAALQRI
DDGTGYCEETGEFGLKRLDAPRIATLSIEAQRHERREKVRDE"
complement(8637..9887)
/gene="AGR_C_2517"
complement(8637..9887)
/gene="AGR_C_2517"
/notes="hypothetical protein"
/codon_start=1
/product="AGR_C_2517p"

Query Match      26.1%; Score 86.2; DB 1; Length 10495;
Best Local Similarity 59.0%; Pred. No. 1.8e-07;
Matches 148; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 80 TCATGGAGCGCGTTCGACCGCGTGGACGATGACCGCTGAGCGTGAACCGC 139
DB 829 TCATGGAGATGCGGTGCGCACTCCGTTCCCGCATGTGATGCGGAATGCGCGCGCCT 888
QY 140 GTGCTGCGCCAGCGCCACGCTTACGTGTCGCCCGAGTGCCTAGATGCGTGAACCGC 199
DB 889 CGCCTGCGCCACCTGCGACGCTGATGTTGACGATGCTGGCAGAGCGTGTGGCGGAC 948
QY 200 CGAGTGACCGAGACGAATGCTGATGTCGATGCGTAGCGGAACGTCGCGCGCATTCGCGGC 259
DB 949 CCGAGCCGATGGAAGAGACATGCTGGCACTTCGCCCTCGAGGTTCTGCCCACTTCGCGGC 1008
QY 260 TGCTCTGCAGATCGCCTTACGACCGCTGCGCGGCTGACCTGCAACTGCGCGGAGG 319
DB 1009 TTTCCTGTCAGATCAGATGATGACGAGCTGACGGTCTCGTTGTCATGTTCCCGAAC 1068
QY 320 CACAGTCATGA 330
DB 1069 GCCAGGCTGA 1079
```

```
RESULT 14
AE009097
LOCUS
DEFINITION
Agrobacterium tumefaciens strain C58 circular chromosome, section
123 of 256 of the complete sequence.
ACCESSION
AE009097 AE008688
VERSION
AE009097.1 GI:17739761
KEYWORDS
Agrobacterium tumefaciens str. C58 (U. Washington)
Agrobacterium tumefaciens str. C58 (U. Washington)
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
1 (bases 1 to 10683)
REFERENCE
AUTHORS
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
QMKRRTTAEALFTWDTLLADIKRHEERHAEIARTHAKSLERQUTSLRQPKSCSEVLQ
```

```
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
The genome of the natural genetic engineer Agrobacterium
tumefaciens C58
Science 294 (5550), 2317-2323 (2001)
21608550
11743193
2 (bases 1 to 10683)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA
Location/Qualifiers
1..10683
/organism="Agrobacterium tumefaciens str. C58 (U.
Washington)"
/mol_type="Genomic DNA"
/strain="C58"
/db_xref="taxon:180835"
209..577
/gene="Atu1349"
209..577
/gene="Atu1349"
/notes="putative; ORF located using Glimmer"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA142355.1"
/db_xref="GI:17739762"
/translation="MAAVSIVFEAPDNYGPAVSKKPTDPEHLAQOTMGDKDLRIEV
LQFRRSAPALHHEMAGADNKVTATAHRLKGSAGVAGSASAAAAVEEKNDLSAL
IAKLAALVFEAEENFILKLCR"
707..1027
/gene="fdx"
/notes="synonym: Atu1350"
707..1027
/gene="fdx"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/product="ferrodoxin"
/protein_id="AA142356.1"
/db_xref="GI:17739763"
/translation="MTKLTIAPDGTPEHLDVSGSTWENAVRSVPCIDACCGAC
ACATCHVYDDAWAERVGPPEWEDMLDFAFEVPTSLSCQIKNNDELGLGVHVP
ERQG"
complement(1334..1960)
/gene="Atu1351"
complement(1334..1960)
/gene="Atu1351"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/product="conserved hypothetical protein"
/protein_id="AA142357.1"
/db_xref="GI:17739764"
/translation="MTKSSALCAVAGALALSLGPAFAETIVKTVSYFVSQKTA
ADDQQLSGHPLTRNTGARHPGATEIKFGSLTYVKMGMSVGTAKVLLTRILLP
QMKRRTTAEALFTWDTLLADIKRHEERHAEIARTHAKSLERQUTSLRQPKSCSEVLQ
```



```

gene
2065..2925
/locus="folf"
/notes="synonym: Atul352"
2065..2925
/locus="folf"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="dihydropterolate synthase"
/protein_id="AAL42358.1"
/db_xref="GI:17739765"
/translation="MTTGSNAQAAGHRSKLDGRIMATVNPATPSFSDGGRYLA
VDAFSAHLCVVEGADIIDIGESRPGAAAVTEAEQDRVLPIEKLRCETDVLIS
VDTYRAATARLAIGAGAHVNDVFGLOKDPDAGVVASTRAGICIMHTGRDRKLDV
IDDPFLNRSLEIAEAGIALDAVLDGFGFKADESNVALMARFGEAAGLPLV
AGTSRKRFGSGTGRVABERDITGAATAILRLAGASIFRVHNVATRDALAIADAV
LAKASAKADA"
3005..3370
/locus="folfB"
/notes="synonym: Atul353"
3005..3370
/locus="folfB"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="dihydroneopterin aldolase"
/protein_id="AAL42359.1"
/db_xref="GI:17739766"
/translation="MSRYTILLKNCSPFAHGLBQEEVLGQFFVDAEMEVEAGDAL
ETDIENTVDIGVAFATKSVVGGRRYLIALNDIAKALRARIPOIWLKLTVRKP
SAPVPGILDYAVQVSEHRA"
3354..3890
/locus="folf"
/notes="synonym: Atul354"
3354..3890
/locus="folf"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="6-hydroxymethyl-7,8-dihydropterin
pyrophosphokinase"
/protein_id="AAL42360.1"
/db_xref="GI:17739767"
/translation="MNIAPSKPTVAALGLGNGTGDPAAMARALRELDHDCRVQAV
SGIYRTPPWKTQADFFNCCALVETSLTALALLQLCLDIERMKKRVTRWGSPRTID
IDVLTGNEAIVTESIEVPHPRTERAFVIMPLADIAPBLEVKGKPVQWLQQAADRS
IVRANKEKRWNTPLGDD"
complement(3905..4465)
/locus="Atul355"
complement(3905..4465)
/locus="Atul355"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42361.1"
/db_xref="GI:17739768"
/translation="MYSRFFSLMGMGLAAILSVSGFGMSDVHAASDRDDVFRSIVAG
SWKGPGEIVAGKYKTKFTGLDTEPLDDKGTGKLGCTCRVGVFQPMVAISQKGN
SYEGKFLDGAAGKGLDITTTQVSGDKVVGINRAKLGAMVAVRSNDNSMNTISYKV
ADOMIPVIGTVTLARDIDQIAVGSIKP"
complement(4595..5674)
/locus="Atul356"
complement(4595..5674)
/locus="Atul356"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1

```

```

/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42362.1"
/db_xref="GI:17739769"
/translation="MKAPTQNDPQTRPAAFLETEEAARPSATQCARPSRDAEISL
TPSDPPLAPADIDAAALFVATPKKSRFSGKGLGALGVLPFLSAPFWADOLIRNL
FSSDVLGTATIALIYALFVAVLGVREVFGIMRLNAVQSLKADETASLDKSPKA
RAIWTIRNLAVLSHRAETAKALKETENDVIGPHLI ELAERELIYLPDRQARAIL
NSSKRVSVTVAVSPRAVVDLAVLFEVTRLVRMAELYGRCPTGLMKLLRVDVAHL
AVTGSIAVGDGLAQVLGHGLASKLARLGEVINGLMTARIGIAAMDLCRPLPFRPV
KRGIGDFMSDLTPDLSGKNKEKA"
complement(5671..7149)
/locus="Atul357"
complement(5671..7149)
/locus="Atul357"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL42363.1"
/db_xref="GI:17739770"
/translation="MPPSLTSLTDSALIAFDNLADRAGNLTHPTLGLVGTLSRAGKT
PDSRLVHNLNGRLPLPFAKSRVSNVRLPQDDAIPRQYEDHIALVDRILW
PDSTRATSEIRITLDYQASGWRWFSAGKLSIDIVDYGWLLDPLLSQDYKQFSD
VTSIAVLSLIRLQSLASLDINADPMTARLAEFAVLKACKSDESLIS
ATVALAGSIRALAEWLAIASSLDINADPMTARLAEFAVLKACKSDESLIS
TLPPRFIMFDDLEGSFALTFAFPGVTDKAKPKSLRAMERRIYAKSIVVKKPRR
EHFARLDQIVLITIQAVNRGPEAVODLERGLDVLCAPRGNTSILSLIRRIDK
VLAAATKADHLHSHDRLERLTRLRITATTGNGAGIEVWALASVATREASVR
QDGHLEPVIPTGPMAGETINGEIDGNKRTAIFGDLPEDEPLFRSIDQDGDKATLP
DVNVRFRPNIDEPGSGGIRLSVPHIRLDRAVQFLFGDKLA"
complement(7278..7787)
/locus="Atul358"
complement(7278..7787)
/locus="Atul358"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

```

Query Match 26.1%; Score 86.2; DB 1; Length 10683;
 Best Local Similarity 59.0%; Pred. No. 1.8e-07;
 Matches 148; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB						
80	TCATGGAGCGCGCTTCTGAAACGGCGTGGACGATATCGAACGCCGTTTGGCGGGGCGCCT	139	777	TCATGGAGATGCGGTGCGCACTCCGTTCCGCGCATTTGATCGGAAATGCGCGGCGCCT	836	140	GTGCTGCGCCACGCTGCCACGCTTTACGTTGCGCCCGAGTGGCTAGATGCGCTGAACCGC	199	837	GCGCTGCGCCACCTGCCACGCTCTATGTTACGATGCTGGCGAGCGCTTGGCGGAC	896	200	CGAGTGAGACCGAAGAACGAAATGCTCGATTGCTGAGCGGAAACGTGCGCGGCGCATTCGCGGC	259	897	CCGAGCGCGATGGAAGAGGACATGCTGCGACTTCGCTTCGAGGTTTCGTCGCCACTTCGCGGC	956	260	TGCTCTGCGCAGATCGCCTTACCGACCTGCTCGACGCGCTGACCCCTGGAACCTGCGGAAGG	319	957	TTTCTGTCGATCAAGATGAATGACGAGCTTGACGCTTCGTTCTGTTTCATGTTCCCGAAC	1016

RESULT 15
 SME591788/c
 LOCUS
 DEFINITION
 Sinorhizobium meliloti 1021 complete chromosome; segment 7/12.
 ACCESSION
 AL591788
 VERSION
 AL591788.1
 GI:15074579
 KEYWORDS
 Sinorhizobium meliloti (Rhizobium meliloti)
 Sinorhizobium meliloti
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
1 (bases 1 to 306250)
Capela,D., Barloy-Hubler,F., Gouzy,J., Bothe,G., Ampe,F., Batut,J.,
Boisard,P., Becker,A., Boutry,M., Cadieu,E., Dreano,S., Gloux,S.,
Grodie,T., Goffeau,A., Kahn,D., Kiss,E., Lelaure,V., Masuy,D.,
Pohl,T., Portetelle,D., Puchner,A., Purnelle,B., Ramsperger,U.,
Renard,C., Thebault,P., Vandenberg,M., Weidner,S. and Galibert,F.
Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021
Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)
21396507
11481430
PUBMED
2
(bases 1 to 306250)
Gouzy,J.
Direct Submission
Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO
EU Consortium
MELILO EU Consortium:
Laboratoire de Biologie Molculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Genetique et Developpement UMR6061-CNRS,
Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,
D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,
Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte
des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,
B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@toulouse.inra.fr
http://sequence.toulouse.inra.fr/meliloti.html.
Location/Qualifiers
1. .306250
/organism="Sinorhizobium meliloti"
/mol_type="genomic DNA"
/strain="1021"
/db_xref="taxon:382"
488..2707
/gene="SMC00954"
488..2707
/gene="SMC00954"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by FrameD"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC46225.1"
/db_xref="GI:15074580"
/db_xref="GOA:Q92PT4"
/db_xref="SPTREMBL:Q92PT4"
/translation="MPAQLHPANAPNIVIMLDDVGPALPDTFGGVIETPTLSRLAQ
DGLSYNFHNAACSPSRAALLTGRNHRVNGQIAELANDWDGYTCRIPRTSATVAK
VLGYGVATAGFGKHYETPDNTNKAVTWKQVHALTPDREFEMYNAPGAAGPHIF
VENTVYDPSHGKEGHTFTEDMTNKAVTWKQVHALTPDREFEMYNAPGAAGPHIF
KWDVAKYKGFDDGDWARRREIYAROKLQWIPONTQLTPTPDTLSGWEDIPDEKXF
ORRLMEVAGTVEADYVQAGRLLEALDELGIRNTLLIFYVWNGNSAQEQNTISEL
LAQGIKTEIKDRIANWELGMDVLGSPKADNMIHAGWAGSTPHRSKLVAAHFG
TRTPLVVSFPGKIPKDPOTRSQPHLNDIPTYDVIDIQPKLVGDISQDLGVS
MTYTFDPSEASGQKEQYFEIGSRATYQDEWIASVFGRIPWAGLCPAIFKWSPDN
DLWELHLSODYSQAKVAAADHPKVBALKRAFQVQAEANKVPVGGGLWSAVFHPED
APSNPATEFOQDVGVGPFTAPKVGARSLNLTIEALRPSDEGVLYALGAFSGGVA
LWVENGKLTVEYNLFEIERTPLESSDPLPSGKNIEVETKVSDDHAPLDVVIRIDG
KEVAKGVPPSPAPLAFTANDAPDVGKDSYFVALYFDRKFPFANIGSKLRVEYLK"
2887..4062
/gene="SMC00955"
2887..4062
/gene="SMC00955"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical

Gene name confidence : hypothetical
predicted by Codon usage
predicted by FrameD"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC46226.1"
/db_xref="GI:15074581"
/db_xref="GOA:Q92PT3"
/db_xref="SPTREMBL:Q92PT3"
/translation="MLLYGCACRPTGVLEPVAANPSSAAQVEMLVTTTSRAKPKGPMF
SGERALSTFAEMTVSIPDPTVRKAGEVAMPKLPNPETDPTATKAVRGGAAEKW
LATHVYKSPDGSVLVFIHGFNNHFDFAVERFAQIIHDSGARSVPVLAWPSRGSLLAY
GYDRETTYRNAVERLPOYLARDPEVKEVSIILAHSGMGNWLSLRCOWAIPNDGLPA
KFRNMLAAPDVVDVDFGQIADMGKQRPRTILFVSRDRDALAVSRVAGVUSRLGAI
DFEQPYKEELAAAGITVDDTKIKAGINDHTKFAESFEIVRLIGSSGQITLDS
RLGLGDHIVAATAGAAHTVGTGAAGLVAAAPALVDQNTRRNYVHVHVALSPSGATR"
4548..4724
/gene="SMC00959"
4548..4724
/gene="SMC00959"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by FrameD"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL/UNKNOWN PROTEIN"
/protein_id="CAC46227.1"
/db_xref="GI:15074582"
/db_xref="SPTREMBL:Q92PT2"
/translation="MSKRELIDTGTDRYVRDSESGKFKHSDVDRSLSSDKKQKAKH
DAKPGQDKGDRKN"
4728..5522
/gene="xthal OR SMC00956"
4728..5522
/gene="xthal OR SMC00956"
/EC_number="3.1.11.2"
/function="macromolecule metabolism; macromolecule
degradation; degradation of dna"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon usage
predicted by Homology
predicted by FrameD"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="PROBABLE EXODEOXYRIBONUCLEASE III PROTEIN"
/protein_id="CAC46228.1"
/db_xref="GI:15074583"
/db_xref="GOA:Q92PT1"
/db_xref="SPTREMBL:Q92PT1"
/translation="MKIATYVNGVGRGLVLLRWLEASPDVVCLOBLKAPDPKFPV
KAIEAGYGAWHGOKSNVGVALLARDREPTLTRKGLQDPPDTHSRVIEAAVEGMVI
GCYLIPNGNPFPGPKFYLAFHRLTAYAAELLELDVPLVILAGDYNVMTLVDYK
ERWNLDFRVEDAYRRLLEQGTIDALRQHPGERVYTTWDFIRNAPFADAGLRID
HLLLSPHVTLLSLAAGVDRVRGEMTSDHAPAWIELSDGPAEEQD"
5519..5737
/gene="SMC00958"
5519..5737
/gene="SMC00958"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by FrameD"
/codon_start=1
/evidence=not experimental
/transl_table=11

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 03:08:33 ; Search time 144.642 Seconds
(without alignments)
9692.234 Million cell updates/sec

Title: US-10-658-691-2
Perfect score: 330
Sequence: 1 atgacgcgaagtcgcgct.....tgcgcgaagcacagtcgcatga 330

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	74	22.4	8766	7 ACC48896	Acc48896 Rhodococc
2	74	22.4	12342	7 ACC48895	Acc48895 Rhodococc
3	74	22.4	23656	7 ACC48894	Acc48894 Rhodococc
C 4	65.6	19.9	306	5 AAH65575	AAH65575 C glutami
5	65.6	19.9	318	5 AAH65577	AAH65577 C glutami
6	65.6	19.9	441	4 AAF71697	AAf71697 Corynebac
C 7	65.6	19.9	34980	5 AAH68525	AAh68525 C glutami
8	62.8	19.0	389	4 ABL07331	ABl07331 Drosophil
9	62.8	19.0	2389	4 ABL07330	ABl07330 Drosophil
10	62.8	19.0	4531	4 ABL07682	ABl07682 Drosophil
11	62.8	19.0	4531	4 ABL07688	ABl07688 Drosophil
12	60.4	18.3	458	8 ACH38677	Ach38677 Human foe
13	60.4	18.3	473	8 ACH33408	Ach33408 Human end
14	60.4	18.3	834	6 ABQ60909	ABq60909 Signal pe
15	60.4	18.3	840	3 AAZ98240	Aaz98240 Human sig
16	60.4	18.3	958	4 AAD16583	Aad16583 Human nov
17	60.4	18.3	958	4 AAD16590	Aad16590 Human nov
18	60	18.2	632	7 ABZ36606	Abz36606 Human GEN
19	60	18.2	1166	7 ABZ36446	Abz36446 Human GEN
20	58.2	17.6	413	7 ABX42738	Abx42738 Bovine ES
21	58.2	17.6	443	7 ABX49100	Abx49100 Bovine ES
22	57.6	17.5	1950	9 ADE73150	Ade73150 Pseudomon
23	57.6	17.5	2232	2 AAQ13693	Aaq13693 Putidared

24	54.4	16.5	300	3 AAC30981	Aac30981 Human sec
25	44	13.3	612	6 ABZ32192	Abz32192 Candida a
26	43.8	13.3	393	7 ACD95816	Acd95816 Human col
27	43	13.0	508	5 AAS64988	Aas64988 DNA encod
28	42.2	12.8	44377	2 AAT80414	Aat80414 Platenoli
29	42.2	12.8	44377	2 AAT80414	Aat80414 Platenoli
30	41.8	12.7	1501	5 AAS91987	Aas91987 DNA encod
31	41.8	12.7	1506	5 AAS92974	Aas92974 DNA encod
C 32	41.8	12.7	10614	4 AAS46259	Aas46259 DNA encod
C 33	41.4	12.5	37856	3 AAA11992	Aaa11992 S. cellul
34	41	12.4	396	6 AAL50587	Aal50587 Bovine ad
35	40.6	12.3	354	6 ABN75809	Abn75809 Human ORF
36	40.2	12.2	3256	5 AAS92977	Aas92977 DNA encod
37	39.6	12.0	3807	7 ACA38183	Aca38183 Prokaryot
38	39.2	11.9	572	2 AAQ43718	Aaq43718 Mel clust
39	39.2	11.9	1442	2 AAQ92327	Aaq92327 Chloropla
40	39.2	11.9	1883	3 AAA28414	Aaa28414 ORF-438-I
41	39.2	11.9	4294	2 AAQ92325	Aaq92325 Plasmid p
C 42	38.8	11.8	2000	7 ADA71938	Ada71938 Rice gene
C 43	38.6	11.7	113193	7 AAD54645	Aad54645 Streptomy
C 44	38.4	11.6	110000	2 AAT42063_03	Continuation (4 of
45	38	11.5	426	7 ABZ40263	Abz40263 N. gonorr

ALIGNMENTS

RESULT 1
ACC48896
ID ACC48896 standard; DNA; 8766 BP.

XX ACC48896;

DT 11-AUG-2003 (first entry)

DE Rhodococcus ruber eth gene cluster involved in ETBE degradation.
XX Cytochrome P450; ethA; ethB; ethC; ethD; ethR; fuel; ether;
KW ethyl tert-butyl ether; ETBE; degradation; bioremediation;
KW soil decontamination; pollutant; biosensor; gene; ds.
XX Rhodococcus ruber.

XX Key Location/Qualifiers
FT CDS 461..4858
FT /tag= a
FT /product= "EthR"
FT /note= "the CDS is claimed in Claim 2, encodes ABR42004"
FT CDS 1640..2878
FT /tag= b
FT /product= "EthA"
FT /note= "the CDS is claimed in Claim 2, encodes ABR42000"
FT CDS 2948..4150
FT /tag= c
FT /product= "EthB"
FT /note= "the CDS is claimed in Claim 2, encodes ABR42001"
FT CDS 4156..4476
FT /tag= d
FT /product= "EthC"
FT /note= "the CDS is claimed in Claim 2, encodes ABR42002"
FT CDS 4547..4858
FT /tag= e
FT /product= "EthD"
FT /note= "the CDS is claimed in Claim 2, encodes ABR42003"

EP1270722-A1.

02-JAN-2003.

22-JUN-2001; 2001EP-00401657.

22-JUN-2001; 2001EP-00401667.

XX


```
Query Match 22.4%; Score 74; DB 7; Length 12342;
Best Local Similarity 54.0%; Pred. No. 6.2e-10;
Matches 174; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 2 TGACCGCAAGGTCGGCGTGTATCTTCGGCGACGCGCGGCTTCGAGCATCTGGTCGAAA 61
Db 4148 TGATGCGGATGCCCAAGATCACCTTCTCCCAATCGGACGGTGTGTCGATCACCGTGTG 4207
QY 62 CCGAGCGGGAGTATCGCTCATGGAAGCGCGCTTCTGAACGGGTGGAGCGTATCGAAG 121
Db 4208 CGTGTGTCGACGAGCGTATGACGCGCGCGCTTCGCGCGAGGTATCGACGGCATCTCG 4267
QY 122 CCGTTTCGGCGGGCGCTGTGCTTGCCTGCGCCACGTCGACGTTTACGTTGGCGCGAGTGGC 181
Db 4268 CCGAGTGTGGTGGCAACCAAGTGTTCCTCCACCTGCCACGTTAGTTCGAAACCGGACAC 4327
QY 182 TAGATGCGCTGAACCGCGGAGTGAGACCGAAGACGAAATGCTCGATG---CGTAGCGG 238
Db 4328 TCGGGCTCTCGCGGACCTGAGCGCGCGAAGACGACATGCTCGACTGCGCGAGCGCG 4387
QY 239 AACGTGCGCGCATTCGCGCTGTCTGCGCAGATCGCGCTTACCGACCTGCTCGACGGCC 298
Db 4388 AACGTGCGAGCAACGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4447
QY 299 TGACCCCTGGAACCTGCCAAGGC 320
Db 4448 TCGGCTGGAGTCCCGACGC 4469

RESULT 3
ACC48894
ID ACC48894 standard; DNA; 23656 BP.
AC ACC48894;
XX
XX
XX 11-AUG-2003 (first entry)
XX Rhodococcus ruber eth gene region involved in ETBE degradation.
XX Cytochrome P450; ethA; ethB; ethC; ethD; ethR; fuel; ether;
XX ethyl tert-butyl ether; ETBE; degradation; bioremediation;
XX soil decontamination; pollutant; biosensor; gene; ds.
XX Rhodococcus ruber.
XX
XX Key Location/Qualifiers
FH repeat_region 3519..23436
FT /tag=c
FT /rpt_type=FLANKING
FT repeat_unit 3515..9094
FT /tag=d
FT /note="duplicated sequence, referred to in Claim 4"
FT misc_feature 9095..21436
FT /tag=a
FT /note="region deleted in ETBE-negative bacteria,
FT referred in Claim 3"
FT misc_feature 9095..17860
FT /tag=b
FT /note="eth gene cluster region, referred to in Claim 7"
FT CDS 9555..13952
FT /tag=f
FT /product="EthR"
FT /note="the CDS is claimed in Claim 2, encodes ABR42004"
FT CDS 10734..11972
FT /tag=g
FT /product="EthA"
FT /note="the CDS is claimed in Claim 2, encodes ABR42000"
FT CDS 12042..13244
FT /tag=h
FT /product="EthB"
FT /note="the CDS is claimed in Claim 2, encodes ABR42001"
FT CDS 13250..13570
FT /tag=i
```

```
FT /product="EthC"
FT /note="the CDS is claimed in Claim 2, encodes ABR42002"
FT CDS 13641..13952
FT /tag=j
FT /product="EthD"
FT /note="the CDS is claimed in Claim 2, encodes ABR42003"
FT repeat_unit 17861..23536
FT /tag=e
FT /note="duplicated sequence, referred to in Claim 4"
XX
XX
XX EP1270722-A1.
PN 02-JAN-2003.
XX
XX 22-JUN-2001; 2001EP-00401667.
XX
XX 22-JUN-2001; 2001EP-00401667.
XX (INSP) INST PASTEUR.
XX (INSF) INST FRANCAIS DU PETROLE.
XX Chauvaux S, Miras I, Beguin P;
XX
XX WPI; 2003-334761/32.
XX
XX P-PSDB; ABR42000, ABR42001, ABR42002, ABR42003, ABR42004.
XX New Rhodococcus ruber nucleic acid, useful for transforming bacteria for
XX depolluting soil contaminated with ethyl tert-butyl ether comprises the
XX cytochrome P-450 gene cluster involved in cleavage of ethyl tert-butyl
XX ether.
XX Claim 7; Page 15-21; 58pp; English.
XX
XX The present sequence is a BamHI fragment of Rhodococcus ruber strain CNCM
XX I-1889 DNA that includes an eth gene cluster encoding a cytochrome P450
XX system involved in the cleavage of ether fuel additives. Expression of
XX the gene cluster is induced by ethyl tert-butyl ether (ETBE). Loss of
XX ability to degrade ETBE is associated with a chromosomal rearrangement
XX resulting in a 14.3 kb deletion including the eth gene cluster. This
XX rearrangement occurs between two 5.6 kb duplicated sequences. The eth
XX gene cluster comprises the ethA, ethB, ethC, ethD and ethR open reading
XX frames. These encode a transcriptional activator (EthR), a ferredoxin
XX reductase (EthA), an ETBE-inducible cytochrome P450 (EthB), a ferredoxin
XX (EthC) and an ETBE-inducible unknown protein (EthD). Identification of
XX the eth gene cluster allows the construction of probes and biosensors for
XX detecting pollution by an ether fuel and for assessing the potential of a
XX contaminated soil or effluent to cleave the ether fuel. Also provided are
XX recombinant bacteria, plant, fungus or yeast cells comprising a vector
XX encoding the eth genes and which are capable of ETBE (and tert-butyl
XX alcohol) degradation or mineralization, for use in ether fuel
XX bioremediation of a contaminated soil or effluent
XX
XX Sequence 23656 BP; 3867 A; 7590 C; 7936 G; 4263 T; 0 U; 0 Other;
XX
XX Query Match 22.4%; Score 74; DB 7; Length 23656;
XX Best Local Similarity 54.0%; Pred. No. 6.9e-10;
XX Matches 174; Conservative 0; Mismatches 145; Indels 3; Gaps 1;
QY 2 TGACCGCAAGTCCGGTGTATCTTCGGCGACGCGCGGCTTCGAGCATCTGGTCGAAA 61
Db 13242 TGATGCGGATGCCCAAGATCACCTTCTCCCAATCGGACGGTGTGTCGATCACCGTGTG 13301
QY 62 CCGAGCGGGAGTATCGCTCATGGAAGCGCGCGCTTCTGAACGGGTGGAGCGTATCGAAG 121
Db 13302 CGTGTGTCGACCGAGCGTGTGCTGCGCGCGCGCTGCGCGAGGTATCGACGGCATCTCG 13361
QY 122 CCGTTTCGGCGGGCGCTGTGCTTGCCTGCGCCACGTCGCGCACTTTACGTTGGCGCGAGTGGC 181
Db 13362 CCGAGTGTGGTGGCAACCGCACGTGTTCACCTGCGCACGTGTACGTCGAACCGGACAC 13421
QY 182 TAGATGCGCTGAACCGCGCGAGTGAGACCGAAGACGAAATGCTCGATG---CGTAGCGG 238
Db 13422 TCGGGCTCTCTCGGACCTGACGCGCGGAGAGACGACATGTCTGACTGCGCGAGCGCG 13481
```

QY 239 AACGTGCGCGCATTCGCGGTGCTCTGCAGATCGCTTACGACCTGCTCGACGCC 298
DB 13482 AACGTGCGAGCACACCGCGCTGCGGTGCTCTGCAGATCGCTTACGACCTGCGACGCC 13541
QY 299 TGACCTGGAAGTCCGGAAGC 320
DB 13542 TCCGGCTGGAGGTCCCGGACGC 13563

RESULT 4
AAH65575/c
ID AAH65575 standard; DNA; 306 BP.

XX AC AAH65575;
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum coding sequence fragment SEQ ID NO: 610.
XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis; ds.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX XX (KYOW) KYOWA HAKKO KOGYO KK.
XX PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI; 2001-376931/40.
XX DR P-PSDB; AAG90356.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX PS Claim 8; SEQ ID NO 610; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX SQ Sequence 306 BP; 75 A; 92 C; 71 G; 58 T; 0 U; 0 Other;

Query Match 19.9%; Score 65.6; DB 5; Length 306;
Best Local Similarity 58.2%; Pred. No. 5.7e-08;
Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

QY 56 TCGAACCAGCGGAGATCGCTCATGAGCGCGCGCTTGTGACCGCGTGGACGCTA 115
DB 286 TCGAGCGGACTGTTGGTATTCAGTAATGAGACCCAGTCGGAACGGAGTGCTCGAA 227
QY 116 TCGAAGCCGTTTTCGGGGGCGCGTGTGCTGCGCCACGTCGACGTTACGTTGGCCCG 175

DB 226 TTGTTGCTGATGGGCGGTTCCTTATCGTGTGCAACCTGCCATGTTGTTGACCTG 167
QY 176 AGTGGCTAGATGCGCTGAAACCCCGAGTAGACCGAGACGAAATGCT---CGATTGCG 232
DB 166 CACAGTATGATGCGCTTCCCCCAATGGAGAGATGAATGCTGTGCGGTGCTG 107
QY 233 TAGCGGAACGTGGCGCGCATTCGCGGTGCTGCTGCCAGATCCGCCCTTACCGA 284
DB 106 CCGTGGACCGTGAGGATTGCTCCCGTTGTCTTGCCCAATCAAGTCAACCGA 55

RESULT 5
AAH65577
ID AAH65577 standard; DNA; 318 BP.

XX AC AAH65577;
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum coding sequence fragment SEQ ID NO: 612.
XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis; ds.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX XX (KYOW) KYOWA HAKKO KOGYO KK.
XX PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI; 2001-376931/40.
XX DR P-PSDB; AAG90358.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX PS Claim 8; SEQ ID NO 612; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX SQ Sequence 318 BP; 71 A; 75 C; 90 G; 82 T; 0 U; 0 Other;

Query Match 19.9%; Score 65.6; DB 5; Length 318;
Best Local Similarity 58.2%; Pred. No. 5.7e-08;
Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

QY 56 TCGAACCAGCGGAGATCGCTCATGAGCGCGCGCTTGTGACCGCGTGGACGCTA 115
DB 47 TCGAGGCGACTGTTGGTATTCAGTAATGAGACCCAGTCGGAACGGAGTGCTCGAA 106

QY 116 TCGAAGCGTTTCGGGGGCGCTGCTGCTGCGCCACGCTGCGACGTTTACGTTGGCCCG 175
Db 107 TTGTTGCTGAATGCGGCGTTCCTTATCGTGCGAACCTGGCATGTTGTTGACCCGTG 166
QY 176 AGTGGCTAGTAGCTGCTGAACCGCGGAGTGAGACCGAAGACGAAATGCT---CGATTGGC 232
Db 167 CACAGTATGATGGGCTTCCCCCAATGAGAGAGATGGAAGATGAATGCTGTGGGGTGTG 226
QY 233 TAGCGGAACGTGCGCGCATTCGCGCTGTCCTGCGCAGATCCGCCCTTACCGA 284
Db 227 CCGTGGACCGTGAGGATTGCTCCGTTTGTCTTGCCCAATCAAGTCAACCGA 278

RESULT 6

AAAF71697

ID AAF71697 standard; DNA; 441 BP.

XX AAF71697;

AC AAF71697;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:675.

XX Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.

XX Corynebacterium glutamicum.

XX W0200100844-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB0000943.

XX 25-JUN-1999; 99US-0141031P.

PR 08-JUL-1999; 99DE-01031412.

PR 08-JUL-1999; 99DE-01031413.

PR 08-JUL-1999; 99DE-01031419.

PR 08-JUL-1999; 99DE-01031420.

PR 08-JUL-1999; 99DE-01031424.

PR 08-JUL-1999; 99DE-01031428.

PR 08-JUL-1999; 99DE-01031431.

PR 08-JUL-1999; 99DE-01031433.

PR 08-JUL-1999; 99DE-01031434.

PR 08-JUL-1999; 99DE-01031510.

PR 08-JUL-1999; 99DE-01031562.

PR 08-JUL-1999; 99DE-01031634.

PR 09-JUL-1999; 99DE-01032180.

PR 09-JUL-1999; 99DE-01032227.

PR 09-JUL-1999; 99DE-01032230.

PR 09-JUL-1999; 99US-0143208P.

PR 14-JUL-1999; 99DE-01032924.

PR 14-JUL-1999; 99DE-01032973.

PR 14-JUL-1999; 99DE-01033005.

PR 27-AUG-1999; 99DE-01040765.

PR 31-AUG-1999; 99US-0151572P.

PR 03-SEP-1999; 99DE-01042076.

PR 03-SEP-1999; 99DE-01042079.

PR 03-SEP-1999; 99DE-01042086.

PR 03-SEP-1999; 99DE-01042087.

PR 03-SEP-1999; 99DE-01042088.

PR 03-SEP-1999; 99DE-01042095.

PR 03-SEP-1999; 99DE-01042123.

PR 03-SEP-1999; 99DE-01042125.

(BADI) BASF AG.

XX

XX

XX

XX

XX

XX

XX

XX

Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

WPI; 2001-061975/07.

P-P5DB; AAB79580.

XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes.

XX Claim 3; Page 1096; 1246pp; English.

XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for
CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteinogenic or
CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein regions required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)

XX Sequence 441 BP; 109 A; 102 C; 116 G; 114 T; 0 U; 0 Other;

Query Match 19.9%; Score 65.6; DB 4; Length 441;

Best Local Similarity 58.2%; Pred. No. 6.1e-08;

Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

QY 56 TCGAAGCGGAGTATCGCTTCATGGAAGCGCGCTTCTGAACGGCTGGACGGTA 115

Db 147 TCGAGCGACTGTTGGTCAATTCAGTATGAGACCGCAGTCCGAACGAGTGGCTGAA 206

QY 116 TCGAAGCGGTTTGGGGGGCGCTGTGCTGCGGCACGTGCCACGTTTACGTTGCCCG 175

Db 207 TTGTTGCTGAATGCGGGCGGTTCTCTATCGTGTGCAACCTGCATGTTGTTGACCTG 266

QY 176 AGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGCGAAATGCT---CGATTGCG 232

Db 267 CACAGTATGATGCGCTTCCCAATGAGAGATGGAAGTGAATGCTGTGGGGTGTG 326

QY 233 TAGCGGAACGTGCGCGCGCATTCGCGGGTGTCTGCGCAGATCCGCCCTTACCGA 284

Db 327 CCGTGGACCGTGAGGATTGCTCCCGTTTGTCTTGCCCAATCAAGTCAACCGA 378

RESULT 7

AAH68525/c

ID AAH68525 standard; DNA; 349980 BP.

XX AAH68525;

AC AAH68525;

DT 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 7060.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX EF1108790-A2.

XX 20-JUN-2001.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280886.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX
PS Disclosure; SEQ ID NO 7060; 246pp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC corynebacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids, described
CC particularly L-lysine. The present sequence is a nucleic acid, described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP; 80289 A; 91081 C; 97378 G; 81232 T; 0 U; 0 Other;
Query Match 19.9%; Score 65.6; DB 5; Length 349980;
Best Local Similarity 58.2%; Pred. No. 1.9e-07;
Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;
QY 56 TCGAACCAGGAGGAGTATCGCTCATGGAAGCGCGCGCTTCTGAACGCGCGTGGACGGTA 115
Db 262917 TCGAGGCGACTGTGGTGATTAGTAATGAGACCGCAGTCCGAACGAGTGCTGGAA 262858
QY 116 TCGAAGCCGTTTGGGGGCGCTGTGCTGCGCCACGTGCGCACGTTTACGTTGCCCCG 175
Db 262857 TTGTTGCTGAATGCGGCGGTTTCTTATCGTGTGCAACCTGCGCATGTGTTGTGACCCTG 262798
QY 176 AGTGGCTAGATCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCT---CGATTGCG 232
Db 262797 CACAGTATGATGCGCTTCCCCAATGAGAGAGATGGAAGATGAATGCTGTGGGTGCTG 262738
QY 233 TAGCGAAGCGTGGCGCGAATTCGGGCTGTCTGCGCAGATCGGCTTACCGA 284
Db 262737 CGGTGGACCGTGAAGATTGCTCCCGTTTCTTGTGCAAAATCAAGGTACCGA 262686
RESULT 8
ABL07331
ID ABL07331 standard; cDNA; 389 BP.
XX
AC ABL07331;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16475.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PSDB; ABB63228.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX
PS Claim 1; SEQ ID NO 16475; 21pp + Sequence Listing; English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 389 BP; 106 A; 97 C; 107 G; 79 T; 0 U; 0 Other;
Query Match 19.0%; Score 62.8; DB 4; Length 389;
Best Local Similarity 56.2%; Pred. No. 3.3e-07;
Matches 119; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 111 CGTATCGAAGCGGTTTGGGGGCGCCCTGTGCTCGCCGACGTGCCAGTTACGTTGG 170
Db 97 CGAGATGGAAGGCGCCTGTGAGGCTTCGCTGGCGCTGCACCCACCTGTACGTTGACGTCCA 156
QY 171 CCCGAGTGGCTAGATCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCTCGATTG 230
Db 157 GCATGATTACCTGCAGAGTTAAGAGCCGAGGAGCAGGAGGAGCTGCTGATAT 216
QY 231 CGTAGCGGAACGTGCGCCGCAATTCGGGGCTGTCTGCCAGATCCGCTTACGACCTGCT 290
Db 217 GGCGCCATTTCTGGCGGAGAACTCCCGGCTGGCTGTGAGTACTCTCGACAAGAGTAT 276
QY 291 CGACGCGCTGACCCCTGGAATCGCCGAGGC 320
Db 277 GGAGGGCATGGAACTGGAGCTGCCAAGGC 306
RESULT 9
ABL07330
ID ABL07330 standard; cDNA; 2389 BP.
XX
AC ABL07330;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16472.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.


```
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB63227.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 16472; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2389 BP; 702 A; 510 C; 492 G; 685 T; 0 U; 0 Other;
SQ
Query Match 19.0%; Score 62.8; DB 4; Length 2389;
Best Local Similarity 56.2%; Pred. No. 4.6e-07;
Matches 118; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 111 CGGTATCGAACCGCTTTGCGGGGCGCTGTGCTGGCGCACGTCGACGCTTTACCTGG 170
DB 1097 CGAGATGGAAGGCCCTGTGAGGCTTCGTGGCTGCACCACTGTACGTGACGTCCA 1156
QY 171 CCCGAGTGGCTAGATCGCTGAAACCCGCGAGTGAGACCGAAGACGAAATGCTCGATTG 230
DB 1157 GCATGATTACTGCGAGAAGTTAAAGAGGCGGAGGAGGAGGAGCGACCTGCGGATAT 1216
QY 231 CGTAGCGGACGTCGCGCGCATTCGCGGCTGCTGCGCAGATCGGCTTACCGACCTGCT 290
DB 1217 GGCGCCATTCTGCGCGAGAACTCCCGGCTCGGCTGTCAGATCTCTCGACAAGAGTAT 1276
QY 291 CGACGGCTGACCTGGAACCTGCCGAAGGC 320
DB 1277 GGAGGGCATGGAATCGAGCTGCCAAGGC 1306
RESULT 10
ABL07682
XX ABL07682 standard; cDNA; 4531 BP.
XX
XX ABL07682;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17528.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 23-MAR-2000; 2000US-0191637P.
```

```
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB63579.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 17528; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 4531 BP; 1185 A; 994 C; 1043 G; 1309 T; 0 U; 0 Other;
SQ
Query Match 19.0%; Score 62.8; DB 4; Length 4531;
Best Local Similarity 56.2%; Pred. No. 5.1e-07;
Matches 118; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 111 CGGTATCGAACCGCTTTGCGGGGCGCTGTGCTGGCGCACGTCGACGCTTTACCTGG 170
DB 339 CGAGATGGAAGGCCCTGTGAGGCTTCGTGGCTGCACCACTGTACGTGACGTCCA 398
QY 171 CCCGAGTGGCTAGATCGCTGAAACCCGCGAGTGAGACCGAAGACGAAATGCTCGATTG 230
DB 399 GCATGATTACTGCGAGAAGTTAAAGAGGCGGAGGAGGAGGAGGAGGAGGAGTAT 458
QY 231 CGTAGCGGACGTCGCGCGCATTCGCGGCTGCTGCGCAGATCGGCTTACCGACCTGCT 290
DB 459 GGCGCCATTCTGCGCGAGAAGTCCCGGCTCGGCTGTCAGATCTCTCGACAAGAGTAT 518
QY 291 CGACGGCTGACCTGGAACCTGCCGAAGGC 320
DB 519 GGAGGGCATGGAATCGAGCTGCCAAGGC 548
RESULT 11
ABL07688
XX ABL07688 standard; cDNA; 4531 BP.
XX
XX ABL07688;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17546.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
```

PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-658660/75.
XX P-PSDB; ABB63585.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 17546; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB2072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4531 BP; 1185 A; 994 C; 1043 G; 1309 T; 0 U; 0 Other;
Query Match 19.0%; Score 62.8; DB 4; Length 4531;
Best Local Similarity 56.2%; Pred. No. 5.1e-07;
Matches 118; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 111 CGGTATCGAAGCGTTTGGGGGGCGCTGTGCGCCGCGACGTCGACGTTTACGTTGG 170
DB 339 CGAGATGGAAGCGCTGTGAGGCTTCGTGCGCTGCACCACTGTACGTGACGTCCA 398
QY 171 CCCCAGTGGCTAGATGCGCTGAACCCGCGAGTGAGACCGAAGACGAAATGCTCGATTG 230
DB 399 GCATGATTACCTGCAGAGTTAAAGAGCGCGAGGAGGAGGACGACCTGCTGGATAT 458
QY 231 CGTAGCGGAACGTGCGCCGCTTCCGGCTGTCTGCGCAGATCCGCTTACGACCTGCT 290
DB 459 GGGCCATTCTGGCGGAGAACTCCCGCTCGGCTGTCTGATCTCTCGACAGAGTAT 518
QY 291 CGACGCGCTGACCTGGAACTCCCGAAGGC 320
DB 519 GGAGGGCATGGAACGTGGAGCTGCGCAAGGC 548
RESULT 12
ACH38677
ID ACH38677 standard; cDNA; 458 BP.
XX ACH38677;
XX
DT 13-OCT-2003 (first entry)
DE Human foetal brain cDNA #44.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX 17-APR-2003.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA
PA (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 25889; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 458 BP; 90 A; 132 C; 152 G; 79 T; 0 U; 5 Other;
Query Match 18.3%; Score 60.4; DB 8; Length 458;
Best Local Similarity 55.1%; Pred. No. 1.1e-06;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 105 CQTGAGCGGTATCGAAGCGCTTTTGGGGGGCGCTGTGCGCGCGCTGCTGCCACGCTTTA 164
DB 211 CGGGTGGAGCTGGAAGGGGCTGTGAAGCTCTCTGGCTGTCTCCACTGCTGCTGTA 270
QY 165 CGTTGGCCCCGAGTGGCTAGATCGCTGAAACCGCCGAGTGAGACCGAAGACGAAATGCT 224
DB 271 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGGAGGAGACGACATGCT 330
QY 225 CGATTGGGTAGCGAAGCGTGGCGCGCTGCTGCGGCTGCTGCCAGATCGGCTTACCGA 284
DB 331 AGACATGGCCCCCTCTCTCCAGAGAACTCGCGGCTGGGCTGCCAGATGCTGACACC 390
QY 285 CTTGCTCGACGGCTGACCTGGAACCTGGAACCTGCGGAAG 318
DB 391 GGAGCTGGAAGGAGCGGAATTCACCTCGCCCAAG 424
RESULT 13
ACH33408
ID ACH33408 standard; cDNA; 473 BP.
XX ACH33408;
XX
DT 13-OCT-2003 (first entry)
DE Human endothelial cell cDNA #1541.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX 17-APR-2003.
XX

XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX DR WPI; 2003-615964/58.
XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX PS Claim 1; SEQ ID NO 20620; 44pp; English.
XX CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX SQ Sequence 473 BP; 96 A; 146 C; 146 G; 83 T; 0 U; 2 Other;
Query Match 18.3%; Score 60.4; DB 8; Length 473;
Best Local Similarity 55.1%; Pred. NO. 1.5e-06;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 105 CGTGGACGGTATCGAAGCCGTTTTCGGGGCGCGCTGTGCTGCGCCACGTCGACGTTTA 164
Db 151 CGGGGTGGACCTGGAAGGGCGCTGTGAAGCTCTCCCTGCGCTGCTCCACCTGCCATGTGTA 210
QY 165 CGTTGGCCCCGAGTGTAGATCGCTGTAACCGCGAGTGAGACGAGAGCAATGCT 224
Db 211 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGAGGAGACGATGCT 270
QY 225 CGATTGCGTAGCGGAACGTCGCGCGCATTCGGGGCTGTCTGTCAGATCCGCTTACCAG 284
Db 271 AGACATGGCCCCCTCTCCAGAGAACTCGCGGCTGGCTGCCAGATTGTGTCACACC 330
QY 285 CTGTCTGACGGCTGACCTTGGNACTGCCGAAG 318
Db 331 GGAGCTGGAAGGAGCGGAATTACCCCTGCCCAAG 364
RESULT 14
ABQ60909
ID ABQ60909 standard; cDNA; 834 BP.
XX AC ABQ60909;
XX DT 26-FEB-2003 (first entry)
XX DE Signal peptide HSP-132 encoding sequence.

KW Neuroprotective; immunomodulator; cancer; chromosome 1p; cytostatic;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnery; gene; ss.
XX OS Homo sapiens.
XX WO200231111-A2.
XX PN 18-APR-2002.
XX PD 11-OCT-2001; 2001WO-US027760.
XX PF 12-OCT-2000; 2000US-00687527.
XX PR (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-426278/45.
XX DR N-FSDB; ABP43665.
XX PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX PS Claim 1; SEQ ID # 122; 357pp + Sequence Listing; English.
XX CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnery, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammatory. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABQ60788-
CC ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 834 BP; 178 A; 237 C; 285 G; 134 T; 0 U; 0 Other;
Query Match 18.3%; Score 60.4; DB 6; Length 834;
Best Local Similarity 55.1%; Pred. No. 1.7e-06;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 105 CGTGGACGGTATCGAAGCCGTTTTCGGGGCGCGCTGTGCTGCGCCACGTCGACGTTTA 164
Db 289 CGGGGTGGACCTGGAAGGGCGCTGTGAAGCTCTCCCTGCGCTGCCATGTGTA 348
QY 165 CGTTGGCCCCGAGTGTAGATCGCTGTAACCGCGAGTGAGACGAGCAATGCT 224
Db 349 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGAGGAGACGATGCT 408
QY 225 CGATTGCGTAGCGGAACGTCGCGCGCATTCGCGGCTGTCTGTCAGATCCGCTTACCAG 284
Db 409 AGACATGGCCCCCTCTCCAGGAGAACTCGCGCTGGGCTGCCAGATTGTGTCACACC 468
QY 285 CTGTCTGACGGCTGACCTTGGNACTGCCGAAG 318
Db 469 GGAGCTGGAAGGAGCGGAATTACCCCTGCCCAAG 502
RESULT 15
AAZ98240
ID AAZ98240 standard; cDNA; 840 BP.
XX

AC AAZ98240;
XX
DT 11-MAY-2000 (first entry)
XX
DE Human signal peptide containing protein HSPp-132 cDNA SEQ ID NO:266.
XX
KW Human; signal peptide-containing protein; HSPp; diagnosis; cancer;
KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
KW antimicrobial; neutropic; neuroprotective; cardiovascular; hepatotropic;
KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;
KW reproductive disorder; developmental disorder; arteriosclerosis;
KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW Parkinson's disease; Huntington's diseases; ovulatory defect;
KW muscular dystrophy; ss.
XX
OS Homo sapiens.
XX
FN WO200000610-A2.
XX
PD 06-JAN-2000.
XX
PF 25-JUN-1999; 99WO-US014484.
XX
PR 26-JUN-1998; 98US-0090762P.
PR 31-JUL-1998; 98US-0094983P.
PR 01-OCT-1998; 98US-0102686P.
PR 11-DEC-1998; 98US-0112129P.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;
XX
XX WPI: 2000-160673/14.
DR P-PSDB; AAY87355.
XX
XX
PT New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and cardiovascular
PT disease.
XX
PS Claim 9; Page 326; 327pp; English.
XX
XX AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the
CC human signal peptide-containing proteins HSPp-1 to HSPp-134. HSPps have
CC anticancer, anti-inflammatory, antimicrobial, neutropic, hepatotropic,
CC neuroprotective, cardiovascular and antiasthmatic activities, and can be
CC used in gene therapy. HSPps can be used to treat or prevent disorders
CC associated with decreased activity or function of HSPp. Antagonists of
CC HSPp are used to treat or prevent disorders associated with increased
CC activity or function of HSPp. Such diseases include cell proliferation
CC (including cancer), inflammation, cardiovascular, neurological,
CC reproductive or developmental disorders, (e.g. arteriosclerosis,
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC asthma, Crohn's disease, microbial or other infections, congestive or
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPp
CC nucleic acids can be used for the recombinant production of HSPp, for
CC detecting HSPp in standard hybridisation and amplification assays (for
CC diagnosis and monitoring), in gene therapy, as antisense, triplex-forming
CC or ribozyme therapeutics, for detecting related sequences or genetic
CC variations, and for chromosomal mapping. HSPp are also used to raise
CC specific antibodies (Ab) and to screen for agonists and antagonists
CC (potential therapeutic agents). Ab are used to diagnose, or monitor, HSPp
CC -related diseases (in usual immunoassays), as therapeutic antagonists, in
CC competitive drug screens, and for purification of HSPp from natural
CC sources
XX
SQ Sequence 840 BP; 178 A; 241 C; 287 G; 134 T; 0 U; 0 Other;

Query Match 18.3%; Score 60.4; DB 3; Length 840;
Best Local Similarity 55.1%; Pred. No. 1.7e-06;

	Matches	118;	Conservative	0;	Mismatches	96;	Indels	0;	Gaps	0;
Qy	105	CCTGACCGTATCGAAGCGTTTGGGGGGCCCTGTGCTGCGCCACGTGCCACGTTTA	164							
Db	298	CGGGGTGGACCTCGAAGGGGGCCTGTGAAGCCTCCCTGCGCTCCACCTGCCATGTGT	357							
Qy	165	CCTTGGCCCCCGAGTGGCTAGATCGCTGAAACCGCGGAGTGAGACCGAAGCGAAATGCT	224							
Db	358	TGTGAGTGAAGACCCACCTGGATCTCTGCTCTCTCCGAGGAGAGGAGAGCATGCT	417							
Qy	225	CGATTGCGTAGCGGAAACGTGCGCCCGCATTCGCGGCTGTCTCTCCAGATCCGCTTACCGA	284							
Db	418	AGACATGGCCCCCTCTCTCCAGAGAACTCGGGGCTGGGCTGCCAGATTGTCTGCACAC	477							
Qy	285	CCTGCTCGAGCGCCTGACCCCTGGAATGCGCGAAG	318							
Db	478	GGAGCTGGAAGGAGCGGAATTCACCTTGCCCAAG	511							

Search completed: June 19, 2004, 04:29:10
Job time : 149.642 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 04:06:17 ; Search time 963.906 Seconds
(without alignments)
10223.528 Million cell updates/sec

Title: US-10-658-691-2
Perfect score: 330
Sequence: 1 atgaccgcaaaagtcgcgct.....tgccgaaggcacagtcacga 330

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71.4	21.6	615	14 CF851075	CF851075 psMA016x0
2	70.2	21.3	578	28 AQ651128	AQ651128 Sheared D
3	66.4	20.1	917	14 CF241493	CF241493 AGENCOURT
4	65.8	19.9	624	9 AL859568	AL859568 AL859568

5	65.8	19.9	714	13 BX733466	BX733466 BX733466
6	65.8	19.9	762	14 CF343736	CF343736 AGENCOURT
7	65.8	19.9	880	14 CF376289	CF376289 AGENCOURT
8	65.8	19.9	913	14 CF241492	CF241492 AGENCOURT
9	65.6	19.9	525	28 BH195739	BH195739 TC3-33B9.
10	65.6	19.9	667	12 BM643564	BM643564 170006873
11	64.8	19.6	979	14 CF220100	CF220100 AGENCOURT
12	64.4	19.5	483	14 CB405072	CB405072 OSTR036H1
13	64.4	19.5	488	14 CB385967	CB385967 OSTF036H1
14	62.8	18.0	577	12 BI626767	BI626767 RH57819.5
15	61.8	18.7	632	9 AI562398	AI562398 TENS2389
16	61.4	18.6	624	9 AI562390	AI562390 TENS2384
17	60.4	18.3	493	12 BM823886	BM823886 K-EST0095
18	60.4	18.3	500	10 AW248266	AW248266 K-EST0095
19	60.4	18.3	588	10 BE222719	BE222719 hu51h11.x
20	60.4	18.3	596	10 BE883887	BE883887 601510295
21	60.4	18.3	599	12 BM786293	BM786293 K-EST0064
22	60.4	18.3	608	12 BG831648	BG831648 602765726
23	60.4	18.3	630	9 AA569348	AA569348 mh89b09.s
24	60.4	18.3	636	10 BF036698	BF036698 601459786
25	60.4	18.3	639	9 AI453531	AI453531 t57a11.x
26	60.4	18.3	644	12 BM682117	BM682117 UI-E-B01-
27	60.4	18.3	648	12 BG775214	BG775214 602650217
28	60.4	18.3	704	12 BI667562	BI667562 603292392
29	60.4	18.3	707	13 BU502531	BU502531 AGENCOURT
30	60.4	18.3	714	14 CA421861	CA421861 UI-H-EGO-
31	60.4	18.3	716	13 BQ957421	BQ957421 AGENCOURT
32	60.4	18.3	743	12 BI823913	BI823913 603039086
33	60.4	18.3	747	10 BE744826	BE744826 601575055
34	60.4	18.3	768	10 BE731453	BE731453 601567736
35	60.4	18.3	768	12 BI762440	BI762440 603048832
36	60.4	18.3	773	14 CB987987	CB987987 AGENCOURT
37	60.4	18.3	824	12 BG775846	BG775846 602650217
38	60.4	18.3	883	12 BM045595	BM045595 603623754
39	60.4	18.3	894	14 CD557260	CD557260 AGENCOURT
40	60.4	18.3	901	13 BQ958408	BQ958408 AGENCOURT
41	60.4	18.3	1201	9 AL524608	AL524608 AL524608
42	60.2	18.2	524	9 AI670909	AI670909 wa06902.x
43	59.4	18.0	1029	13 BU502474	BU502474 AGENCOURT
44	59.2	17.9	282	9 AA341290	AA341290 EST46665
45	59	17.9	521	9 AA824331	AA824331 aj30d11.s

ALIGNMENTS

RESULT 1	CF851075	psMA016x004f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation_sMA Phytophthora sojae cDNA clone sMA016004 5, mRNA sequence.	615 bp	mRNA	linear	EST 30-OCT-2003
LOCUS	CF851075	CF851075	615 bp	mRNA	linear	EST 30-OCT-2003
DEFINITION	CF851075	CF851075	615 bp	mRNA	linear	EST 30-OCT-2003
ACCESSION	CF851075	CF851075	615 bp	mRNA	linear	EST 30-OCT-2003
VERSION	CF851075.1	GI:38066729	615 bp	mRNA	linear	EST 30-OCT-2003
KEYWORDS	EST.	EST.	615 bp	mRNA	linear	EST 30-OCT-2003
SOURCE	Phytophthora sojae	Phytophthora sojae	615 bp	mRNA	linear	EST 30-OCT-2003
ORGANISM	Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.	Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.	615 bp	mRNA	linear	EST 30-OCT-2003
REFERENCE	1 (bases 1 to 615)	1 (bases 1 to 615)	615 bp	mRNA	linear	EST 30-OCT-2003
AUTHORS	Tyler, B.	Tyler, B.	615 bp	mRNA	linear	EST 30-OCT-2003
TITLE	Tyler, B. Not Published	Tyler, B. Not Published	615 bp	mRNA	linear	EST 30-OCT-2003
JOURNAL	Unpublished (2003)	Unpublished (2003)	615 bp	mRNA	linear	EST 30-OCT-2003
COMMENT	Contact: Tyler B	Contact: Tyler B	615 bp	mRNA	linear	EST 30-OCT-2003
	Tyler lab	Tyler lab	615 bp	mRNA	linear	EST 30-OCT-2003
	VBI	VBI	615 bp	mRNA	linear	EST 30-OCT-2003
	1880 Pratt Dr., Blacksburg, VA 24061, USA	1880 Pratt Dr., Blacksburg, VA 24061, USA	615 bp	mRNA	linear	EST 30-OCT-2003
	Tel: 540-231-7318	Tel: 540-231-7318	615 bp	mRNA	linear	EST 30-OCT-2003
	Email: bmtylev@vt.edu	Email: bmtylev@vt.edu	615 bp	mRNA	linear	EST 30-OCT-2003
	PCR Primers	PCR Primers	615 bp	mRNA	linear	EST 30-OCT-2003
	FORWARD: BK reverse primer	FORWARD: BK reverse primer	615 bp	mRNA	linear	EST 30-OCT-2003
	BACKWARD: BK reverse primer	BACKWARD: BK reverse primer	615 bp	mRNA	linear	EST 30-OCT-2003
	Plate: 016 row: O column: 04	Plate: 016 row: O column: 04	615 bp	mRNA	linear	EST 30-OCT-2003

```

Seq primer: BK reverse primer
High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source           1..615
     /organism="Phytophthora sojae"
     /mol_type="mRNA"
     /db_xref="taxon:67593"
     /clone="sMA016004"
     /tissue_type="mycelium"
     /cell_line="P6437"
     /dev_stage="mycelium"
     /lab_host="synthetic medium"
     /clone_lib="USDA-IFAPS:Expression of Phytophthora sojae
genes during infection and propagation sMA"
     /notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: Xho1"

ORIGIN
Query Match       21.6%; Score 71.4; DB 14; Length 615;
Best Local Similarity 54.0%; Pred. No. 2.8e-05;
Matches 169; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY      8  CAAAGTCCCGGTGATCTTCGGCGCAGCCGGCGCTTCGAGCATCTGGTCGAACCGAAG 67
      |||
Db      69  CGAGGTGACCTCAAGTTCAGCTGCGGACACTCCATCAAGAGGTGGAGGCCAAGA 128
      |||

QY      68  CGGAGTATCGCTCATGGAAGCGCGCTTCTGAACGGCGGTGGACGGTATCGAAGCGTTT 127
      |||
Db      129  CGGGCATGTCCATCTCGACGTGGCCACGCCAACGACATCGAC---CTCGAAGGGCGT 185
      |||

QY      128  GCGGGGCGCGCTGTGCTCGCGCCACGTGCCACGTTTACGTTGGCCCCGAGTGGCTAGATG 187
      |||
Db      186  GCAGTGTCTCATGGCTGTCTCACGTGCACAGTGATCTTGAGGACCCCGTGTTCGACG 245
      |||

QY      188  CGGTGAACCGCGAGTAGACCGAAGACGAATGTCGATTCGTAGCGGAACGTCGCC 247
      |||
Db      246  AGCTGAGGAGGCGCTGCGAGGACGAGGACATGCTGGACATGGCGCTTCGCGCTCACGC 305
      |||

QY      248  CGCATTCGCGGCTGCTGTCAGATCCGCTTACGACCTGCTCGACGCGCTGACCTGG 307
      |||
Db      306  ACACGTCAGATTAGGCTGCAGGTGTTCTGTGACGAGGGCTTCGAGGGCACGCCGTGA 365
      |||

QY      308  AATGCGGAAGGC 320
      |||
Db      366  CGCTGCCCAAGGC 378
      |||

RESULT 2
AQ651128      578 bp DNA linear GSS 22-JUN-1999
LOCUS          Sheared DNA-20N12.TF Sheared DNA Trypanosoma brucei genomic clone
DEFINITION     Sheared DNA-20N12, genomic survey sequence.
ACCESSION      AQ651128
VERSION        AQ651128.1 GI:5144314
KEYWORDS       GSS.
SOURCE         Trypanosoma brucei
ORGANISM       Trypanosoma brucei
               Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
               Trypanosoma.
REFERENCE      1 (bases 1 to 578)
AUTHORS        El-Sayed,N., Zhao,S., Zhao,H., Gili,S., Suh,E., Malek,J., Fujii,C.,
               Gerrard,C., Leach,V., de Jong,P., Ullu,E., Melville,S.,
               Donelson,J., Fraser,C. and Adams,M.
               Determination of clone end sequences from Trypanosoma brucei GUTat
               10.1 sheared DNA library
               Unpublished (1999)
COMMENT        Other GSSs: Sheared DNA-20N12.TF
               Contact: Najib M. El-Sayed
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0200
               Fax: 301 838 0208
               Email: nelsayed@tigr.org

Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ARCC. Sheared DNA end sequences search page:
http://www.tigr.org/tdb/mdb/tbdb/.
Seq primer: M13-Forward
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..578
     /organism="Trypanosoma brucei"
     /mol_type="genomic DNA"
     /strain="TREU927/4 GUTat 10.1"
     /db_xref="taxon:5691"
     /clone="Sheared DNA-20N12"
     /clone_lib="Sheared DNA-20N12"
     /note="Vector: pUC18; Site 1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. in Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Barrell, Oxford University
Press, 1999)."
```

ORIGIN

```

Query Match       21.3%; Score 70.2; DB 28; Length 578;
Best Local Similarity 51.6%; Pred. No. 5e-05;
Matches 159; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY      14  TCCGCGTGATCTTCGCGCAGCCGGCGCTTCGAGCATCTGGTCGAACCGAAGCGGAG 73
      |||
Db      193  TCCGCGTCAATGTGACACCGCTGAGGGGAAAGATTACATTCTCTGCCCAAGTGTT 252
      |||

QY      74  TATCGCTCATGGAAGCGCGCTTCTGAACGCGTGGACGGTATCGAAGCGGTTTGGCGGG 133
      |||
Db      253  TAAATTGATGGAGCGCTTCGTGACGTTGCTCGGTGGACATTGAAGCCGCTTGCACG 312
      |||

QY      134  CGGCTGTGCTCGGCCACGTCGTCACGTTTACGTTGGCCCCGAGTGGCTAGATGCGGTGA 193
      |||
Db      313  GGACATCGCCTGTAGCATTGTTCACGTCATCTCTTCGGAAGAGGACTTTGGGAAACTGA 372
      |||

QY      194  AACGCGGAGTGAGACCGAAGACGAATGTCGATTGCGTAGCGGAACGTCGCGCGCAT 253
      |||
Db      373  CAGCGGATCGGAGATGAATGACATGTGGACCTCGCGCCGANGTGAATCCGACGT 432
      |||

QY      254  CGCGCTGTCTGCCAGATCGGCTTACCGACCTGCTCGACGCGCTGACCTCGAACTGC 313
      |||
Db      433  CGCGCTCGGTCGCCAGTAAAGCTCAGCAAGAACTTGATGGGATTACGCTGCAGATGC 492
      |||

QY      314  CGAAGGCA 321
      |||
Db      493  CGAGTGAA 500
      |||
```

RESULT 3

```

CF241493/c
LOCUS          CF241493
DEFINITION     AGENCOURT 15113448 NICHED_XGC_Emb6 Silurana tropicalis cDNA clone
IMAGE:6992976 3', mRNA sequence.
ACCESSION      CF241493
VERSION        CF241493.1 GI:33444701
KEYWORDS       EST.
SOURCE         Silurana tropicalis (western clawed frog)
ORGANISM       Silurana tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
               Xenopodinae; Silurana.
               1 (bases 1 to 917)
REFERENCE      1
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
```

COMMENT

Contact: Daniela S. Gethard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Sequencing by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14672 row: k column: 23
High quality sequence start: 8
High quality sequence stop: 752.
Location/Qualifiers
1. .917
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6992976"
/tissue_type="neurula"
/dev_stages="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHX XGC Emb6"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: this is a
Xenopus Gene Collection (XGC) library."

ORIGIN

Query Match 20.1%; Score 66.4; DB 14; Length 917;
Best Local Similarity 60.2%; Pred. No. 0.00037;
Matches 109; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 116 TCGAACCGTTTGGCGGGCGCTGCTGGCGCCAGTGCAGACGAAATGCTCGATTGCGTAG 175
DB 800 TGAAGGCGCTGTGAATCTCTGCTGCTCCACGTCGACGTTTACGTGAATCTG 741
QY 176 AGTGGCTAGATGCGCTGAAACCGCGGAGTGAGACGGAAGTCTCGATTGCGTAG 235
DB 740 AGTTTTCGATAAATCTCCAGAGCCGATGAAGGAGGATGATGCTGACATGSCAC 681
QY 236 CGGAACGTGGCGGATTCGGCGCTGCTGCGCAGATCCGCTTACGACCTGTCGACG 295
DB 680 GTTTCCTTCAGGAGAACTCCGCGCTCGCTGCGCAGATCATCTCACCAGAACTGAACG 621
QY 296 G 296
DB 620 G 620

RESULT 4

AL859568
LOCUS
DEFINITION AL859568 XGC-egg Silurana tropicalis cDNA clone TEGG060p07 5', mRNA
sequence.
ACCESSION AL859568
VERSION AL859568.2 GI:38655003
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 624)
REFERENCE
Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT On Sep 15, 2002 this sequence version replaced gi:22979749.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TEGG060p07.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1. .624
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGG060p07"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

FEATURES

source

ORIGIN

Query Match 19.9%; Score 65.8; DB 9; Length 624;
Best Local Similarity 60.2%; Pred. No. 0.00045;
Matches 109; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 116 TCGAACCGTTTGGCGGGCGCTGCTGGCGCCAGTGCAGACGTTTACGTGGCCCG 175
DB 260 TGAAGGCGCTGTGAATCTCTGCTGCTCCACGTCGACGTTTACGTGAATCTG 319
QY 176 AGTGGCTAGATGCGCTGAAACCGCGGAGTGAGACGGAAGTCTCGATTGCGTAG 235
DB 320 AGTTTTCGATAAATCTCCAGAGCCGATGAAGGAGGATGATGCTGACATGSCAC 379
QY 236 CGGAACGTGGCGGATTCGGCGCTGCTGCGCAGATCCGCTTACGACCTGTCGACG 295
DB 380 CGTTGCTTCAGGAGAACTCCGCGCTCGCTGCGCAGATCATCTCACCAGAACTGAACG 439
QY 296 G 296
DB 440 G 440

RESULT 5

EX733466
LOCUS
DEFINITION EX733466 XGC-tadpole Silurana tropicalis cDNA clone TTPA053b02 5',
mRNA sequence.
ACCESSION EX733466
VERSION EX733466.1 GI:38406207
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 714)
REFERENCE
Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TTPA053b02.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.

CDNA was oligo dT primed from Sug of poly A⁺ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: *Escherichia coli* DH10B.

[illegible]

```

1. 714
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="UTPA053502"
/dev_host="tadpole (stage 35-4)"
/dev_host="E. coli DH10"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site: 1.
was oligo dt primed from 5'ug of
embryos. EcoRI-NotI cut cDNA was
ligated into EcoRI at the 5' end and
NotI at the 3' end and ligated into

```

ORIGIN

Query Match	19.9%;	Score	65.8;	DB	13;	Length	714;
Best Local Similarity	60.2%;	Pred. No.	0.00047;				
Matches	109;	Conservative	0;	Mismatches	72;	Indels	0;
Gaps	0;						

QY	116	TCGAAGCCGTTTGCGGGGGCCCTGTGCTCGCGCCACGTGCCACGTTTACGTTGGCCCGC	175
Db	439	TGGAAGGCGCCTGTGAATCCTCCCTGGCCTGCTCCACGTGCCACGTTTACGTTGAATCTG	498
QY	176	ATGTCGTAGTCGGCTGAAACCCGCGAGTGAGACCGAAGACGAAATGCTTCGATTCGGTAG	235
Db	499	AGTTTTTCGATAAACTTCACAGGCCGATGAAAGGGAGGATGATATCTGGACATGGCAC	558
QY	236	CGGAACGTGCGCCGATTCGGGGCTGTCTCGCCAGATCCGCTTACCGACCTGCTCCGACG	295
Db	559	CGTTGCTTCAGAGAACTCCCGGCTCGGCTGCCAGATCATTTCCACCGAAGAACTGAACG	618
QY	296	G	296
Db	619	G	619

RESULT 6	CF343736	762 bp	mRNA	linear	EST 18-AUG-2003
LOCUS					
DEFINITION	AGENCOURT_15225932 NICHXD XGC Swb1				
	IMAGE:6995635 5'. mRNA sequence.				

REFERENCE	1. (bases 1 to 762)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgabps-remail.nih.gov Tissue Procurement: Rob Granger, University of Virginia cDNA Library Preparation: Open Biosystems cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAMA14690 row: a column: 10 High quality sequence stop: 681. Location/Qualifiers
FEATURES	

```

source
i. .762
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:699635"
/tissue_type="whole body"
/clone_lib="NICHID_XGC Swbi"
/vector="pExpress-1, Site 1: EcoRV; Site 2: NotI;
Bulk tissue was collected from a whole 10 month old male
from the F8 strain. 1st strand cDNA was primed with a Not
I - oligo(dT) primer, double-stranded cDNA was cloned into
the Not I and EcoRV sites of pExpress-1. Library was
size-selected for >1.5 kb fragments for an average insert
size of 1.92 kb. A normalized version of this library is
also available (NICHID_XGC Swb1N). Library was constructed
by Open Biosystems (Huntsville, AL)."

```

ORIGIN

	Query Match	19.9%;	Score 65.8;	DB 14;	Length 762;
	Best Local Similarity	60.2%;	Pred. No. 0.00048;		
	Matches 109;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;
Qy	116	TCGAACGCGTTTCGGGGGGCGCTGTGCTCGCCACGTCGCCACGTTACGTTGGCCCG	175		
Db	500	TGGAAGCGCGCTGTGAATCCTCCCTGCTGCTCCACGTGCCACGTTACGTGAATCTG	559		
Qy	176	ATGGGCTAGATGGCTTGAAACCGCCGAGTGAGACCGAAGACGAAATGCTCGATTGGGTAG	235		
Db	560	AGTTTTTCGATAACTTCAGAGAGCCCGATGAAGGGAGGATGATGCTGGACATGGCAC	619		
Qy	236	CGGAACGTGGCGCGATTTCGGGCTGCTCTGCCAGATCCGCGTTACCGACCTGCTCGACG	295		
Db	620	CGTGTGTTTCAGAGAACTCCCGGCTCGGTGCCAGATCATCTTCACCGAAGAACTGNAAG	679		
Qy	296	G 296			
Db	680	G 680			

RESULT 7	CF376289	LOCUS	CF376289	880 bp	mRNA	linear	EST 27-AUG-2003
DEFINITION	AGENCOURT_15350861	NICHD_XGC_Swb1N	Siurana tropicalis cDNA clone				
	IMAGE:7004712	5'	mRNA sequence.				

REFERENCE	1 (bases 1 to 880)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: coapbs-xemail.nih.gov Tissue Procurement: Rob Granger, University of Virginia cDNA Library Preparation: Open Biosystems cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LHAM14702 row: h column: 22 High quality sequence stop: 661.

source

```
1. .880
/organism="Silurana tropicalis"
```



```

/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7004712"
/tissue_type="whole body"
/clone_lib="NICHD XGC Swb1N"
/note="Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; Bulk tissue was collected from a whole 10 month old male from the P6 strain. 1st strand cDNA was primed with a Not I - oligo(dT) primer. double-stranded cDNA was cloned into the Not I and EcoRV sites of pExpress-1. Library was size-selected for >1.5 kb fragments for an average insert size of 1.92 kb. Library was normalized to Cots with a 180-fold reduction of actin. A non-normalized version of this library is also available (NICHD XGC Swb1). Library was constructed by Open Biosystems (Huntsville, AL)."

ORIGIN
Query Match 19.9%; Score 65.8; DB 14; Length 980;
Best Local Similarity 60.2%; Pred. No. 0.0005;
Matches 109; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 116 TCGAAGCCGTTTGGGGGGCCCTGTGCTGCGCCACGTCGTCACGTTTACGTTGGCCCCG 175
Db 458 TCGAAGGGCGCTGTGAATCCTCCCTGGCTGCTCCACGTCGTCACGTTTACGTTGACTG 517

QY 176 AGTGCTAGATGCGCTGAAACCGCGGAGTGAGACCGAAGACGAAATGCTGATGCGTAG 235
Db 518 AGTTTTCGATAAACTTCACAGCCCGATGAAGGAGGATGATGCTGGACATGGCAC 577

QY 236 CGGAACGTGCGCGGCTGCTGCTCCAGATCCGCTTACCGACCTGCTCGACG 295
Db 578 CGTTGCTCAGAGAACTCCCGGCTGCGCTGCCAGATCATCTCCCGAAGAACTGAACG 637

QY 296 G 296
Db 638 G 638

RESULT 8
CF241492
LOCUS
DEFINITION
AGENCOURT 15113421 NICHD XGC Emb6 Silurana tropicalis cDNA clone
IMAGE:6992976 5', mRNA sequence.
CF241492
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 913)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcgpb-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution by: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM14672 row: k column: 23
High quality sequence stop: 712.
Location/Qualifiers
1. 913
/mol_type="mRNA"
/organism="Silurana tropicalis"

FEATURES
source

```

```

/db_xref="taxon:8364"
/clone="IMAGE:6992976"
/tissue_type="neural"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb6"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 19.9%; Score 65.8; DB 14; Length 913;
Best Local Similarity 60.2%; Pred. No. 0.0005;
Matches 109; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 116 TCGAAGCCGTTTGGGGGGCCCTGTGCTGCGCCACGTCGTCACGTTTACGTTGGCCCCG 175
Db 458 TCGAAGGGCGCTGTGAATCCTCCCTGGCTGCTCCACGTCGTCACGTTTACGTTGACTG 517

QY 176 AGTGCTAGATGCGCTGAAACCGCGGAGTGAGACCGAAGACGAAATGCTGATGCGTAG 235
Db 518 AGTTTTCGATAAACTTCACAGCCCGATGAAGGAGGATGATGCTGGACATGGCAC 577

QY 236 CGGAACGTGCGCGGCTGCTGCTCCAGATCCGCTTACCGACCTGCTCGACG 295
Db 578 CGTTGCTCAGAGAACTCCCGGCTGCGCTGCCAGATCATCTCCCGAAGAACTGAACG 637

QY 296 G 296
Db 638 G 638

RESULT 9
BH195739/c
LOCUS
DEFINITION
TC3-33B9.TV TC3 Trypanosoma cruzi genomic clone TC3-33B9, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
GSS.
Trypanosoma cruzi
Trypanosoma cruzi
Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae;
1 (bases 1 to 525)
Trypanosoma; Schizotrypanum.
Klug, S., Edwards, K.E., Nilsson, D., Bontempi, E.J., Myler, P.,
Stuart, K., Ghedin, E., El-Sayed, N.M. and Andersson, B.
Clustering and analysis of BAC-end and GSS sequences from
Trypanosoma cruzi
Unpublished (2001)
Contact: Bjorn Andersson
Department of Genetics and Pathology
Uppsala University
Rudbeck Laboratory, SE-751 85, Uppsala, Sweden
Tel: 46 18 471 4107
Fax: 46 18 471 4808
Email: bjorn.andersson@genpat.uu.se
Clones are derived from the Trypanosoma cruzi CL-Brener BAC library
TC3. For clone availability, please contact Dr. Bjorn Andersson at
Uppsala University (bjorn.andersson@genpat.uu.se).
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. 525
/organism="Trypanosoma cruzi"
/mol_type="genomic DNA"
/strain="CL Brener"
/db_xref="taxon:5693"
/clone="TC3-33B9"
/clone_lib="TC3"
/note="Vector: pBelobAC11; Site 1: Hin dIII. Constructed
for Uppsala University by Marie-Christine Le Paslier in

```

the laboratory of Denis Le Paslier at the Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France. Briefly, Trypanosoma cruzi CL-Brener agarose embedded DNA (obtained from Dr. Franco da Silveira) was partially digested with Hin dIII. High molecular weight fragments were ligated in pGEM-T Easy digested with Hin dIII. The average insert size is 100 kb. Total clone coverage: approx. 33 X the haploid genome."

ORIGIN

Query Match 19.9%; Score 65.6; DB 28; Length 525;
Best Local Similarity 54.0%; Pred. No. 0.00047;
Matches 134; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 68 CGGAGTATCGTCTATGGAAGCGCGCTTCTGAAGCGGTGACGGTATCGAAGCGCTT 127
|||
DB 508 CGGGCATTAACCTCATGAGGCAATACGGGACTTGGGAAGACTTGACATTGAGCGCGAT 449
|||

QY 128 GCGGGGCGCTGCTGCGCCAGCTGCCAGCTTTACGTTGCGCCGAGTGGCTAGATG 187
|||
DB 448 GTGATGAACTTGTGCTCGAGCACCCTGCCAGCTGATGATCTCCGAGGAGGACTTTGAGA 389
|||

QY 188 CGTGAAACCGCGAGTGAACGAGAGAAATGCTCGATTGCGTAGCGGAACGTGGCG 247
|||
DB 388 AGCTGAGTGAGCCCTCAGAAGACGAGTGGACATGCTGACCTTGGCGCATCCGTCACAA 329
|||

QY 248 CGCATTCGGGCTGCTCTCCAGATCGGCTTACGACCTGTCGACGCGCTGACCTGG 307
|||
DB 328 AGACATCAGCATATCGTGCCAAATTCGCTTACGAGCGCGCTGATGATTTACCGTGA 269
|||

QY 308 AACTGCGG 315
|||
DB 268 AGTTGCGG 261
|||

RESULT 10

BM643564 667 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687312521 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
DEFINITION 19600449629190 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

CELERA GENOMICS

45 W. Gude Dr.,

Rockville, MD 20850, USA

Tel: 2404533151

Fax: 2404534580

Email: HoltRA@celera.com

Plate: NU01004ABS row: C column: 12

Seq primer: M13 Reverse.

Location/Qualifiers

1. .667

/organism="Anopheles gambiae"

/mol_type="mRNA"

/strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)"

/db_xref="taxon:7165"

/clone="19600449629190"

/dev_stage="Adult"

/lab_host="DH10b"

/clone_lib="A.Gam.ad.cDNA1"

/note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole

FEATURES

source

1. .667

/organism="Anopheles gambiae"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="IMAGE:6988704"

/dev_stage="embryo, stages 10-13"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICH-XGC Embs"

/note="Vector: PCMV-SF0RT5.1; Site_1: NotI; Site_2: EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.0 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."

adult mosquitoes (mixed sex) frozen on liquid nitrogen. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."

ORIGIN

Query Match 19.9%; Score 65.6; DB 12; Length 667;
Best Local Similarity 56.5%; Pred. No. 0.00051;
Matches 122; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 105 CGTGACGGTATCGAAGCCGTTTGGGGGGCGCTGCTGCGCCACGTCGACCGTTTA 164
|||
DB 337 CGGTGTCGATGAGAGGGCGCTGCGAGGCTCGCTGGCTGCGACCTGTCACTGTA 396
|||

QY 165 CGTTGGCCCGAGTGGCTAGATCGCTGAAACCGCCGAGTGAGACGGAACGAAATGCT 224
|||
DB 397 CGTGACGACGATATCTCGATCGTCTGCGCGAGCCGGAGGAGGAGGACGATCTGCT 456
|||

QY 225 CGATTGCTAGCGAAGTGGCGGCTTTCGGGCTGCTGCGAGATCGGCTTACCGA 284
|||
DB 457 CGATATGSCACCGTTTCTGCGCGAAACTCCCGCTCGGCTGCGCAGATAGTGTGAGAA 516
|||

QY 285 CTTGCTCGACGGCTGACCTCGAACTTGGCGAAGGC 320
|||
DB 517 GGACCTCGAAGGATGCGGTTGCAGCTTCCGAGGC 552
|||

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

CELERA GENOMICS

45 W. Gude Dr.,

Rockville, MD 20850, USA

Tel: 2404533151

Fax: 2404534580

Email: HoltRA@celera.com

Plate: NU01004ABS row: C column: 12

Seq primer: M13 Reverse.

Location/Qualifiers

1. .979

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="IMAGE:6988704"

/dev_stage="embryo, stages 10-13"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICH-XGC Embs"

/note="Vector: PCMV-SF0RT5.1; Site_1: NotI; Site_2: EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.0 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."

FEATURES

source

1. .979

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="IMAGE:6988704"

/dev_stage="embryo, stages 10-13"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICH-XGC Embs"

/note="Vector: PCMV-SF0RT5.1; Site_1: NotI; Site_2: EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.0 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

Query Match 19.6%; Score 64.8; DB 14; Length 979;
 Best Local Similarity 59.7%; Pred. No. 0.00084;
 Matches 108; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 116 TCGAAGCCGTTTGGGGGGCCCTGTGCTCGCCACGTCGACGTTTACGTTTGGCCCCG 175
 DB 555 TGAAGGGCCGCTGTGAATCCTCCCTGGGCTGCTCCAGCTGCCAGTTTACGTGAATACTG 614

QY 176 AGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCTCGATTGCGTAG 235
 DB 615 AGTTTTTCGATAAATCTCCAGAGCCCGATGAAGAGGAGATGATATGCTGACATGGCAC 674

QY 236 CGGAACGTCGCGCCGATTCGGGGGTGTCTCTGCCAGATCCGCTTACCGACCTGCTCGACG 295
 DB 675 GTTGTCTTCANGAAGAACTCCCGGCTCGGCTGCCAGATCATTTCCACCGAAGACTGAACG 734

QY 296 G 296
 DB 735 G 735

RESULT 12

CB405072/c
 LOCUS CB405072 483 bp mRNA linear EST 15-MAY-2003
 DEFINITION OSTRO36H11_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
 ACCESSION CB405072
 VERSION CB405072.1 GI:30746799
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE
 AUTHORS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 483)

Reboul, J., Vaglio, P., Rual, J.P., Lamesch, P., Martinez, M.,
 Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T.,
 Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S.,
 Endress, G.A., Jenna, S., Chevet, E., Papasotiropoulos, V.,
 Tolias, P.P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H.,
 Doucette-Stamm, L., Hill, D.E. and Vidal, M.

TITLE C. elegans ORFeome version 1.1: experimental verification of the
 genome annotation and resource for proteome-scale protein
 expression

JOURNAL

COMMENT Nat. Genet. (2003) In press

Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739

Email: Marc.Vidal@dfci.harvard.edu
 Sequence tag of Gateway entry clones. The primers used were
 designed on the predicted protein encoding ORF. C. elegans ORFeome
 cloning project : Contact david_hill@dfci.harvard.edu or
 marc_vidal@dfci.harvard.edu
 POLYA=No.

FEATURES

source

Location/Qualifiers

1. 483
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /sex="Hermaphrodite and male"
 /tissue_type="whole animal"
 /dev_stage="mixed stage"
 /clone_lib="AD-wrmcDNA"
 /note="The AD-wrmcDNA library was generated with poly(A)+
 RNA isolated from both hermaphrodite and male N2 worms of
 all larval stages, embryos, adults and dauers and the
 subsequent generation of cDNAs by poly(A) priming. The
 cDNAs were cloned into pPC86"

ORIGIN

Query Match 19.5%; Score 64.4; DB 14; Length 483;
 Best Local Similarity 57.4%; Pred. No. 0.00084;
 Matches 116; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 115 ATGAAGCCGTTTGGGGGGCCCTGTGCTCGCCACGTCGACGTTTACGTTTGGCCCC 174
 DB 226 ATGAGGGAGCTTGGAGGCATCACTTGGCTCTCGAGTGCACGCTTACGTCGATCCA 167

QY 175 GAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCTCGATTGCGTA 234
 DB 166 GCGTTCAGAACCAAGCTCCCGAGCGCGCTAGAGAGGAGACGATATGCTGGATATGGCT 107

QY 235 GCGGAACGTCGCGCCGATTCGGGGGTGTCTCTGCCAGATCCGCTTACCGACCTGCTCGAC 294
 DB 106 CCGCGCTCAAGGACAACTCACGGCTCGGCTGCTCAATTGTGCTCACGAAGAGCTCGAC 47

QY 295 GCGCTGACCTGGAACCTGCCGA 316
 DB 46 GGAATCACTGTAACCCCTGCCGA 25

RESULT 13

CB385967
 LOCUS CB385967 488 bp mRNA linear EST 15-MAY-2003
 DEFINITION OSTRO36H11_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
 ACCESSION CB385967
 VERSION CB385967.1 GI:30727677
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE
 AUTHORS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 488)

Reboul, J., Vaglio, P., Rual, J.P., Lamesch, P., Martinez, M.,
 Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T.,
 Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S.,
 Endress, G.A., Jenna, S., Chevet, E., Papasotiropoulos, V.,
 Tolias, P.P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H.,
 Doucette-Stamm, L., Hill, D.E. and Vidal, M.

TITLE C. elegans ORFeome version 1.1: experimental verification of the
 genome annotation and resource for proteome-scale protein
 expression

JOURNAL

COMMENT Nat. Genet. (2003) In press

Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739

Email: Marc.Vidal@dfci.harvard.edu
 Sequence tag of Gateway entry clones. The primers used were
 designed on the predicted protein encoding ORF. C. elegans ORFeome
 cloning project : Contact david_hill@dfci.harvard.edu or
 marc_vidal@dfci.harvard.edu
 POLYA=No.

FEATURES

source

Location/Qualifiers

1. 488
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /sex="Hermaphrodite and male"
 /tissue_type="whole animal"
 /dev_stage="mixed stage"
 /clone_lib="AD-wrmcDNA"
 /note="The AD-wrmcDNA library was generated with poly(A)+
 RNA isolated from both hermaphrodite and male N2 worms of
 all larval stages, embryos, adults and dauers and the
 subsequent generation of cDNAs by poly(A) priming. The
 cDNAs were cloned into pPC86"

```
Query Match      19.5%; Score 64.4; DB 14; Length 488;
Best Local Similarity 57.4%; Pred. No. 0.0084;
Matches 116; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 115 ATCGAGCCGTTTGGCGGCGCCTGTGCTGCGCCACGTCACAGTTTACGTTTGGCCCC 174
DB 239 ATGAGGGAGCTTGGCAGGATCCTACTTGGCTGCTGACGTGCACGTCTACGTGATCA 298
QY 175 GAGTGTAGATCGCTGAACCGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 234
DB 299 GCCTTCCAGAACAGCTCCGGAGCGCTAGAGAGAGAGAGAGAGATGCTGATATGGCT 358
QY 235 GCGAGAGCTGCGCGGATTCGGGGCTGCTGCGAGATCGCGCTTACCGACCTGCGAC 294
DB 359 CCGCGCTCAAGGACAACTCACGCTCGGCTGTCAAATTTGCTGTACGAAAGAGCTCGAC 418
QY 295 GCGCTGACCTGGAATGCGGA 316
DB 419 GGAATCACTGTACCTGCGGA 440

RESULT 14
BI626767
LOCUS
DEFINITION
  BI626767 577 bp mRNA linear EST 07-SEP-2001
  RH67819.5prime RH Drosophila melanogaster normalized Head pFlc-1
  Drosophila melanogaster cDNA clone RH67819 5 similar to Fdxh;
  FBan0004205 GO: [ferredoxin (GO:0008042); electron transfer
  (GO:0005489)] located on: 3L 67B1-67B1; 08/24/2001, mRNA sequence.
ACCESSION
  BI626767
VERSION
  BI626767.1 GI:15522292
KEYWORDS
  EST.
SOURCE
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 577)
  Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,
  Carlson,J., Champe,M., Chavez,M., Dorsett,V., Farfan,D., Friese,E.,
  George,R., Gonzalez,R., Guatin,H., Harris,N., Li,P., Liao,G.,
  Misra,S., Mungall,C.J., Munoo,J., Pacleb,J., Paragas,V., Park,S.,
  Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and
  Rubin,G.M.
  BDGP/HMI RH Drosophila EST Project
  Unpublished (2001)
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AB003552; acc:3L [9109492,9396275]
  estimated-cyto:67A1-67B5; 08/24/2001
  Plate: RH.678 row: B column: 7
  High quality sequence stop: 379.
FEATURES
  source
  1..577
  /organism="Drosophila melanogaster"
  /mol_type="mRNA"
  /db_xref="taxon:7227"
  /clone="RH67819"
  /sex="male and female"
  /dev_stage="Adult"
  /lab_host="DHS-alpha Tona"
  /clone_lib="RH Drosophila melanogaster normalized Head
  pFlc-1"
  /note="Organ: head; Vector: pFlc1; Site 1: XhoI; Site 2:
  BamHI; Library was kindly generated by Piero Carninci at
  the RIKEN. The library was normalized and excised using
  Cre recombinase. Plasmid cDNA library."
ORIGIN
  Query Match      19.0%; Score 62.8; DB 12; Length 577;
  Best Local Similarity 57.4%; Pred. No. 0.0084;
  Matches 116; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 115 ATCGAGCCGTTTGGCGGCGCCTGTGCTGCGCCACGTCACAGTTTACGTTTGGCCCC 174
DB 239 ATGAGGGAGCTTGGCAGGATCCTACTTGGCTGCTGACGTGCACGTCTACGTGATCA 298
QY 175 GAGTGTAGATCGCTGAACCGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 234
DB 299 GCCTTCCAGAACAGCTCCGGAGCGCTAGAGAGAGAGAGAGATGCTGATATGGCT 358
QY 235 GCGAGAGCTGCGCGGATTCGGGGCTGCTGCGAGATCGCGCTTACCGACCTGCGAC 294
DB 359 CCGCGCTCAAGGACAACTCACGCTCGGCTGTCAAATTTGCTGTACGAAAGAGCTCGAC 418
QY 295 GCGCTGACCTGGAATGCGGA 316
DB 419 GGAATCACTGTACCTGCGGA 440

RESULT 15
BI626398
LOCUS
DEFINITION
  BI626398 632 bp mRNA linear EST 25-MAR-1999
  TENS2389 T. cruzi epimastigote normalized cDNA library Trypanosoma
  cruzi cDNA clone 2389 5', mRNA sequence.
ACCESSION
  BI626398
VERSION
  BI626398.1 GI:4513743
KEYWORDS
  EST.
SOURCE
  Trypanosoma cruzi
  Trypanosoma cruzi
  Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
  Trypanosoma; Schizotrypanum.
REFERENCE
  1 (bases 1 to 632)
  Verdun,R.E., Di Paolo,N.C., Urmenyi,T.P., Rondinelli,E.,
  Frasch,A.C.C. and Sanchez,D.O.
  Gene discovery through expressed sequence tag sequencing in
  Trypanosoma cruzi
  Infect. Immun. 66 (11), 5393-5398 (1998)
  99003155
  9784549
  Contact: Sanchez D.O.
  Instituto de Investigaciones Biologicas (Univ. Nac. de Gral
  San Martin)
  Av. Gral Paz entre Albarillos y Constituyentes, INTI edificio 24
  cp(1650) San Martin, Prov. de BS AS, Argentina
  Tel: (54-1)752-9639 or (54-1)752-0021
  Fax: (54-1)752-0021 or (54-1)752-9639
  Email: dsanchez@inti.gov.ar
  Seq primer: T7.
FEATURES
  source
  Location/Qualifiers
  1..632
  /organism="Trypanosoma cruzi"
  /mol_type="mRNA"
  /strain="Cl-Brenner"
  /db_xref="taxon:5693"
  /clone="2389"
  /cell_type="epimastigote"
  /clone_lib="T. cruzi epimastigote normalized cDNA library"
  /note="cDNA library constructed with oligo dr primed
  epimastigote mRNA and cloned in pT7318d phagemid with
  modified polylinker (PHARMACIA)"
ORIGIN
  Query Match      18.7%; Score 61.8; DB 9; Length 632;
  Best Local Similarity 56.7%; Pred. No. 0.0033;
  Matches 114; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 115 ATCGAGCCGTTTGGCGGCGCCTGTGCTGCGCCACGTCACAGTTTACGTTTGGCCCC 174
DB 239 ATGAGGGAGCTTGGCAGGATCCTACTTGGCTGCTGACGTGCACGTCTACGTGATCA 298
QY 175 GAGTGTAGATCGCTGAACCGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 234
DB 299 GCCTTCCAGAACAGCTCCGGAGCGCTAGAGAGAGAGAGATGCTGATATGGCT 358
QY 235 GCGAGAGCTGCGCGGATTCGGGGCTGCTGCGAGATCGCGCTTACCGACCTGCGAC 294
DB 359 CCGCGCTCAAGGACAACTCACGCTCGGCTGTCAAATTTGCTGTACGAAAGAGCTCGAC 418
QY 295 GCGCTGACCTGGAATGCGGA 316
DB 419 GGAATCACTGTACCTGCGGA 440
```

Db	110	GAGGACTTTGAGAAAGCTGAGTGAGCCCTCAGAGACGAGGTGGACATGCTCGACTTGGCG	169
Qy	235	GCGGNAAGTGGCGCGCATTCGGGCTGTCTGCCAGATCCGCTTACCGACTGCTCGAC	294
Db	170	CCATCCGTCACAAAGACATCAGATTATCGTGCCAAATTCAGCTTACGGACGGCGCTCGAT	229
Qy	295	GCCCTGACCCCTGGAACTGCCG	315
Db	230	GGTATTACCGTGAAGTTGCCG	250

Search completed: June 19, 2004, 07:19:52
 Job time : 971.906 secs

Db 1431 GGGTCAGCCTGATGACAGGCTGAGTCTCCAAATGGTATCTACGATATTGTGCGTGAATTGT 1490
QY 130 GGGGGCGCTGTGCTTGGCCGACGTCGCCAGTTTACGTTGGCCCGAGTGGCTAGATCGG 189
Db 1491 GGGGGCAGCGCCAGCTGTGCCACCTGCTATGTGTAACGAAGCGTTACGGACAAG 1550
QY 190 CTGAACCGCCGAGTGAGACCGAAGCAATGCTCGATTGGT---AGCGGAAGTGGC 246
Db 1551 GTGCCGCGCCCAACGAGCGGGAATCGGCATGCTGGAGTGGTCAACGCCGAATGGA 1610
QY 247 CCGCAATTCGCGCTGTCTTCCGAGATCCGCTTACCGACCTGCTCGACGGGCTGACCCCTG 306
Db 1611 CCGAACAGCAGGCTGTCTTCCGAGATCATCATGACCGCGAGTGGATCGTGTGTC 1670
QY 307 GAACGCTCGAAGGACAGTCA 330
Db 1671 GATGTTCCCGATAGGCAATGGTAA 1694

RESULT 2
US-09-252-991A-6336
; Sequence 6336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6336
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6336

Query Match 16.4%; Score 54; DB 4; Length 345;
Best Local Similarity 58.3%; Pred. No. 1.4e-05;
Matches 134; Conservative 0; Mismatches 90; Indels 6; Gaps 2;
QY 56 TCGAACCAGGAGGATGCTCATGGAAGCGCGCTTCTGAACGCGCTGGACGTA 115
Db 62 TCGAGCGGAAGCCGCGCAACCATCTGATGCGCGCTGGCAATGGCATCGA---GA 118
QY 116 TCGAAGCCGTTTGGGGGCGCTGTGCTGCGCCACGTCGCCAGCTTTACGTTGGCCCG 175
Db 119 TCGAGCATGCTGTGAGAAATCTGCGCTGCAACCTGCGACCTGCGCAATTCGT---CCGCG 175
QY 176 ATGGCTAGATGCTGAAACCGCGAGTGAGACCGAAGCAAAATGCTCGATTGCGTAG 235
Db 176 AAGGCTCAGTCCATGAGCGCTCGACGAACTCGAAGACGACATGCTGGACAGGCT 235
QY 236 CGGAAGCTGCGCGATTGCGGCTGTCTGCGCATCGGCTTACCGAC 285
Db 236 GGGTCTGAGCCGATTGCGGCTATCTGCGCATCTGCGAGCAGTGGTGGCCGAC 285

RESULT 3
US-09-252-991A-6047/c
; Sequence 6047, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6047
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6047
Query Match 16.4%; Score 54; DB 4; Length 756;
Best Local Similarity 58.3%; Pred. No. 1.6e-05;
Matches 134; Conservative 0; Mismatches 90; Indels 6; Gaps 2;
QY 56 TCGAACCAGGAGGAGTATGCTCATGGAAGCGCGCTTCTGAACGCGCTGGACGTA 115
Db 388 TCGAGCGAAGCCGCGCAACCATCTGATGCGCGCTTGGCAATGGCATCGA---GA 332
QY 116 TCGAAGCCGTTTGGGGGCGCTGTGCTGCGCCACGTCGCCAGCTTACGTTGGCCCG 175
Db 331 TCGAGCATGCTGTGAGAAATCTGCGCTGCAACCTGCGACCTGCGCAATTCGT---CCGCG 275
QY 176 AGTGGCTAGATGCTGAAACCGCGAGTGAGACCGAAGCAAAATGCTCGATTGCGTAG 235
Db 274 AAGGCTCGACTCCATGGAGCGCTCGGCAACTCGAAGACGACATGCTGGCAAGGCT 215
QY 236 CGGAAGCTGCGCGCATTCGCGCTCTGCGCAGATCCGCTTACCGAC 285
Db 214 GGGTCTGGAGCCGATTCGCGCTATCTCTGCCAGCAGTGGTGGCCGAC 165
RESULT 4
US-09-489-039A-4374
; Sequence 4374, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4374
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4374
Query Match 14.6%; Score 48.2; DB 4; Length 363;
Best Local Similarity 55.6%; Pred. No. 0.00048;
Matches 114; Conservative 0; Mismatches 88; Indels 3; Gaps 1;
QY 81 CATGGAAGCGCGCTTCTGAACGCGCTGGAAGCGTATCGAAGCGTTTGGGGGCGCGCTG 140
Db 105 CATCTCGATGCGCGCTGCTGAGCGTATCGAAATTGAACGCTGCGAATAATCTCTG 164
QY 141 TGCTGCGCCACGTCGACGTTTACGTTGGCCCGAGTGGCTAGATGCGCTGAAACCGCC 200
Db 165 TGCTGCAACCACTGCGAC---TGCATCGTGCCTGAAGTTTGTGCTGCGCGAAG 221
QY 201 GAGTGAGACCGAAGCAATGCTGATTCGCTAGCGGAACGTCGCGCGCATTCGCGGCT 260
Db 222 TAGCGAAGATGAAGACGACATGCTGGATAAAGCTGGGCGCTGGAACCTGACACGCTCT 281
QY 261 GTCTGCGCATCCGCTTACCGAC 285
Db 282 GAGTGGCAGCGCGGTCAACCGAC 306

RESULT 5
US-09-252-991A-6212/C
; Sequence 6212; Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6212
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6212

Query Match 13.0%; Score 43; DB 4; Length 2112;
Best Local Similarity 57.6%; Pred. No. 0.015;
Matches 117; Conservative 0; Mismatches 80; Indels 6; Gaps 2;

QY 56 TCGAACCAGAGCGGAGTATCGCTCATGGAAGCGGCCGTTCTGAACGGCGTGGACGTA 115
DB 197 TCGAGGCGAAGCGCGGGAACCATCTGATGCGCGTTGCGCAATGGCATCGA---GA 141

QY 116 TCGAAGCCGTTTGGGGGGCGCTGTGCTGCGCCAGTGCACGTTTACGTTGGCCCG 175
DB 140 TCGAGCATGCTGTGAAATCTTGGCCCTGCGACCTGCCACGTAATGCT---CGCG 84

QY 176 AGTGCGTAGTGGCTGAAACCGCCGAGTGGAGACCGAAGACGAAATGCTGATTCGCTAG 235
DB 83 AAGGCTCGACTCCATCGAGCGCGTCCGACGAACTCGAAGACGACATGCTGGACAAGCCT 24

QY 236 CGGACGTCGCCCGCATTCGCG 258
DB 23 GGGGCTGGAGCGCGATTCGCG 1

RESULT 6
US-09-252-991A-6417
; Sequence 6417; Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6417
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6417

Query Match 12.8%; Score 42.2; DB 4; Length 405;
Best Local Similarity 59.7%; Pred. No. 0.019;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 167 TTGCCCCCGAGTGGCTAGATCGCTGAAACCGCCGAGTGAGACCGAAGACGAAATGCTCG 226
DB 1 TCGTCCCGAAGGCTCGACTCCATGAGCGCTGCGACGAACTCGAAGACGACATGCTGG 60

QY 227 ATTGCTAGCGAAGCTGCGCGCATTCGGGGTGTCTCTGCCAGATCGCGCTTACCGAC 285
DB 61 ACAAGGCTGSGGTCTGGAGCCCATTCGGGCTATCTGCCAGGACGATGCTGGCGGAC 119

RESULT 7
US-08-804-227C-7
; Sequence 7; Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kunstock, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-227C-7

Query Match 12.8%; Score 42.2; DB 2; Length 44377;
Best Local Similarity 47.5%; Pred. No. 0.042;
Matches 125; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 3 GACCCGAAAGTTCGCGGTGATCTCCGCGCAGCGCGGCTTCGAGCATCTGTCGAAAC 62
DB 29022 GCCCGGCGAGGTGCGCGCTCGCGCTAGCGCGCGCGGCTCACTTCGCGGACGCTCAT 29081

QY 63 CGAAGCGGAGGTATCGCTCATGGAAGCGCGCGCTTCTGAACGCGCTGGACGATTCGAAAGC 122
DB 29082 CGCCCTCGCATGTACCCGCGCGAGCGAGATGGAAACGGAGGCGCGGCGGCGCATCGTGT 29141

QY 123 CGTTTGGGGGGCGCTGTGCTGCGCCACGTCGACGTTTACGTTGTCGCCCGCGTGGCT 182
Db 29142 CGAGGTGCGCCCGCGGTACCGGTGTCGCGTCCGCGACCGCGGTGCTCGGCCTGTGGGA 29201
QY 183 AGATCGCTGAAACCGCGAGTGAGACGGAAGACGAAATGCTCGATTGCTGAGCGGAACG 242
Db 29202 CGCGGGCGTGGCGCGCTGTGCGTGGCGGACGACCGCGTGTCTGCGCCCGTCCCGGACG 29261
QY 243 TCGCGCGGCTTGGCGGTGTCT 265
Db 29262 CTGGTCTACGCCCGCGCGCT 29284

RESULT 8
US-08-804-198-1
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostek, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
US-08-804-198-1

Query Match 12.8%; Score 42.2; DB 2; Length 44377;

Best Local Similarity 47.5%; Pred. No. 0.042;
Matches 125; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
QY 3 GACCGAAAGTCCGGTGAATCTTCGCCACGCGCGGCTTCGAGCATCTCGTGAAGC 62
Db 29022 GCCCGCCAGGTGCGGTGCGGTACGCGCCGCGCGCTCAACTTCGCGGACGCCCTCAT 29081
QY 63 CGAAGCGGGAGTATCGCTCATGGAAGCGCGCTTCTGAACGGCGTGGACGGTATCGAAGC 122
Db 29082 CGCCCTCGGCATGTACCCGGGCGAGCGGAGATGGGAACGAGGGCGCGGACCGTCT 29141
QY 123 CGTTTCGGGGCGCTGTGCTTCCTGCGCCACGTCGACCGTTTACGTTGGCCCCGAGTGGCT 182
Db 29142 CGAGGTGCGCCCGCGGTCAACGGTGTGCGCGGACCGCGTCTCGGCTGTGGGA 29201
QY 183 AGATCGCTGAACCCCGAGTGAGACCGAAGACGAAATGCTCGATTGCTAGCGGAACG 242
Db 29202 CGCGGCCTGGGCCCGCTGTGCTGTCGCGGACCGACCGGCTGTGCGCCCGTCCCGACG 29261
QY 243 TCGCGCGCATTCGCGCTGTCT 265
Db 29262 CTGGTCTACGCCCGCGCGCT 29284

RESULT 9
US-09-427-700-8
; Sequence 8, Application US/09427700
; Patent No. 6372489
; GENERAL INFORMATION:
; APPLICANT: AntiCancer
; APPLICANT: Zhao Ming
; TITLE OF INVENTION: METHOD AND MODEL FOR HAIR PIGMENTATION
; FILE REFERENCE: 31276-20017.00
; CURRENT APPLICATION NUMBER: US/09/427,700
; CURRENT FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: US 60/105,725
; PRIOR FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(450)
; OTHER INFORMATION: ORF 438 gene
US-09-427-700-8

Query Match 11.9%; Score 39.2; DB 4; Length 459;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 4 ACCGAAAGTCCGGTGAATCTTCGCCACGCGCGGCTTCGAGCATCTCGTGAAGC 63
Db 121 ACCCCGAGGTCCCGGGAACCGCGCGGTTCGCGCCCGCCCTTCGACGAGATC 180
QY 64 GAAGCGGGAGTATCGCTCATGGAAGCGGCGCTTCTGAACGGGTGGACGGTATCGAAGC 123
Db 181 TACAGGGCGCGCGGATACAGGGCCCGGACGGTACCGACGCGGGCGCCACACGCGCGC 240
QY 124 GTTTGCGGGGCGCTGTGCTTCGCGCCACGTCGACGTTTACGTTGGCCCCGAGTGGCTA 183
Db 241 GGTCACGGCGGTGACGGTCAACGGCGCGGCCCATCACGGCGGGGTTACGCGGTTCGTG 300
QY 184 GATGCGCTGAACCGC 199
Db 301 GACGGGTGCACTGC 316

RESULT 10
US-07-989-363-1
; Sequence 1, Application US/07989363

Patent No. 5385841
GENERAL INFORMATION:
APPLICANT: Ortega, Jose Daza
APPLICANT: Gill, Jose Antonio
APPLICANT: Garcia, Tomas Vigil
APPLICANT: Martin, Juan Francisco
TITLE OF INVENTION: P0438, A NEW CALCIUM-REGULATED PROMOTER
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Washington
STATE: D. C.
COUNTRY: U. S.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/989,363
FILING DATE: 19921211
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1768-112A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 572 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 94..531
US-07-989-363-1

Query Match 11.9%; Score 39.2; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 4 ACCGCAAGGTCGCGGTGATCTTCGCGCAGCGCGCGGCTTCGAGCATCTGGTCGAAACC 63
Db 202 ACCCGCGGTCCCGGGAACCCGCGCGCTTCGCGCGCCCGCGCGCTTCGAGCATC 261
Qy 64 GAAGCGGAGTATCGCTCATGGAAGCGCGCGCTTCGAAAGCGGCTGAGCGGTATCGAAGCC 123
Db 262 TACAAGGCGCGCGCGGATACAGGGCGCGGACCGGTACCGACGCGCGCGCGCGCGC 321
Qy 124 GTTTGCGGGGCGCGCTGTCCTGCGCAGCGTGCACGTTTACGTTGCGCGCGCGCGCTA 183
Db 322 GGTACGCGGCTGACGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 381
Qy 184 GATGCGTGAACCGC 199
Db 382 GACGCGCTCGAATGTC 397

RESULT 11
US-08-264-526-1
Sequence 1, Application US/08264526
Patent No. 566099
GENERAL INFORMATION:
APPLICANT: Ortega, Antonio Daza
APPLICANT: Gill, Jose Antonio
APPLICANT: Garcia, Tomas Vigil

APPLICANT: Martin, Juan Francisco
TITLE OF INVENTION: P0438, A NEW CALCIUM-REGULATED PROMOTER
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 701-E Columbia Square
STREET: 555 13th Street, N. W.
CITY: Washington
STATE: D. C.
COUNTRY: U. S.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,526
FILING DATE: 23-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/989,363
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1768-119A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 572 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 94..531
US-08-264-526-1

Query Match 11.9%; Score 39.2; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 4 ACCGCAAGGTCGCGGTGATCTTCGCGCAGCGCGCGGCTTCGAGCATCTGGTCGAAACC 63
Db 202 ACCCGCGGTCCCGGGAACCCGCGCGCTTCGCGCGCCCGCGCGCTTCGAGCATC 261
Qy 64 GAAGCGGAGTATCGCTCATGGAAGCGCGCGCTTCGAAAGCGGCTGAGCGGTATCGAAGCC 123
Db 262 TACAAGGCGCGCGGATACAGGGCGCGGACCGGTACCGACGCGCGCGCGCGCGCGC 321
Qy 124 GTTTGCGGGGCGCGCTGTCCTGCGCAGCGTGCACGTTTACGTTGCGCGCGCGCGCTA 183
Db 322 GGTACGCGGCTGACGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 381
Qy 184 GATGCGTGAACCGC 199
Db 382 GACGCGCTCGAATGTC 397

RESULT 12
US-08-152-483B-8
Sequence 8, Application US/08152483B
Patent No. 5529909
GENERAL INFORMATION:
APPLICANT: Della-Cioppa, Guy
APPLICANT: Kumagai, Monto
TITLE OF INVENTION: TYROSINASE-ACTIVATOR
TITLE OF INVENTION: PROTEIN FUSION ENZYME
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 2730 Sand Hill Road
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,483B
FILING DATE: No. 5529909ember 12, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7/857,602
FILING DATE: March 30, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 923,692
FILING DATE: July 31, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 600,244
FILING DATE: October 22, 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 641,617
FILING DATE: January 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 737,899
FILING DATE: July 26, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: BIOG-20240/81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-3660
TELEFAX: (415) 854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1442
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM:
IMMEDIATE SOURCE:
CLONE:
FEATURE:
US-08-152-483B-8

Query Match	11.9%;	Score 39.2;	DB 1;	Length 1442;
Best Local Similarity	50.0%;	Pred. No. 0.14;		
Matches	98;	Conservative	0;	Mismatches 98; Indels 0; Gaps 0;

QY	4	ACCGCAAAAGGTC	CGGTGATCTTCGGCGAGCGGGGCTTCGAGCATCTGTCGAACC	63
Db	288	ACCCCGAGGTCCCGGNA	CCGGCGCTTCGGGCGCCCGCCCTTCGACGAGATC	347
QY	64	GAAGCGGGAGTATCGCTCATGAAAGCGCGCGTTCTGAACGGGTGACGGTATCGAAGCC	123	
Db	348	TACAAGGGCGCGGATACAGGGCGGACGGTCACCGACGGCGGGGCCACACGGCGGC	407	
QY	124	GTTCGCGGGGCGCTGTGCCCTGTGCCACGTGCGACGTTTACGTGGCCCCGAGTGGCTA	183	
Db	408	GGTACGCGCGGTGACGFTACGGCGGCGGGCCATCACGGGGGGGTTACGCGGTGTTCTGTG	467	

Query Match 11.9%; Score 39.2; DB 1; Length 4009;

TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE:
 DESCRIPTION: DNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM:
 IMMEDIATE SOURCE:
 CLONE:
 FEATURE:
 US-08-152-483B-2

Db	3336	ACCCCGAGGTC	CCCGGAA	CCCGCC	CGTC	CGGCGCC	CTTC	GACGAGATC	3395
Qy	64	GAAGCGGAGTAT	CGCTCAT	GAAAGCGGCGCTT	TGAACGGCGTG	GAACGGT	TGGAAGCC	123	
Db	3396	TACAAAGGCG	CCGGATAC	AGGGCGGAC	GGTACACG	CGGGGGCC	CACACGCGCG	3455	
Qy	124	GTTCGGGGGG	CGCTGT	CGCTG	CGCCACG	TGTCACG	TTTACGTTGGCCCCGAGTGCGCTA	183	
Db	3456	GGTACGCGGT	GACGTC	ACGCGCGG	CGCCAT	CACGGCGGCGGTTAC	CGCCGTGTCGTG	3515	
Qy	184	GATCGCTGA	AACCGC	199					
Db	3516	GACGGCGTC	GAATCG	3531					

RESULT 15
 US-09-557-884-1/C
 ; Sequence 1, Application US/09557884
 ; Patent No. 6506381
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleischmann et al.
 ; TITLE OF INVENTION: The Nucleotide sequence of
 ; the Haemophilus influenzae Rd Genome. Fragments
 ; Thereof, and Uses Thereof
 ;

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,984
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

Best Local Similarity 50.0%; Pred.No. 0.17;
Matches 98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 4 ACGCCAAAGTCCGCGTGATCTTCCGCGCAGCGCGGGTTCGAGCATCTGGTCGAAACC 63
Db 2526 ACCCCCGAGTCCCGGGAACCCGCGCGGTTCGCGCGCCCCCGCCCTTCGACGAGATC 2585
QY 64 GAAGCGGGAGTATCGCTCATGGAAGCGCGCGTTCGAAACGGCGTGACGGTATCGAAGCC 123
Db 2586 TCAAGGGCGCGGATACAGGGCCGGAACGGTACCGACGGCGGGGGCCACCAACGGCGGC 2645
QY 124 GTTTCGGGGGGCGCTGTGCTCTCGGCCACGTCGCGCAGCTTACGTTGGCCCCGAGTCGCTA 183
Db 2646 GGTACGCGGTGAACGTCACGCGCGGGCCATCAGCGCGGGCGGTACGCGCGGTTCGCTG 2705
QY 184 GATGCGTGAACCGC 199
Db 2706 GACGGGTCGAACTGC 2721

RESULT 14
US-08-152-483B-2
; Sequence 2, Application US/08152483B
; Patent No. 5529909
; GENERAL INFORMATION:
; APPLICANT: Della-Cioppa, Guy
; APPLICANT: Kumagai, Monto
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
; TITLE OF INVENTION: PROTEIN FUSION ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,483B
; FILING DATE: No. 5529909ember 12, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7/657,602
; FILING DATE: March 30, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 923,692
; FILING DATE: July 31, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 600,244
; FILING DATE: October 22, 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 641,617
; FILING DATE: January 16, 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 737,899
; FILING DATE: July 26, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-3660
; TELEFAX: (415) 854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4294

```

```
;
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      11.6%; Score 38.4; DB 4; Length 1830121;
Best Local Similarity 51.3%; Pred. No. 0.81;
Matches 115; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

QY      46 GAGCATCTCGTCAAAACCGAAGCGGAGTATCGCTCATGGAAGCGGCCGCTTCTGAACGGC 105
Db      394522 GAAGGTATGCTGCTTGATGCTGCAACAGGAGATAACTTATTAGAAGTTGCTCACAATGCG 394463

QY      106 GTGGACGGTATCGAAGCCGTTTGGCGGGCGCTGTGCTGGCCACGTCGACGTTTAC 165
Db      394462 GGAGTAGAAATTCTACACCTTTGTATGTTCTTTGTGCTTGACAACTTGCCACGT---G 394406

QY      166 GTTGCCCCCGAGTGGCTAGATCGCTGAAACCGCCGAGTGAGACCGAAGACGAAATGCTC 225
Db      394405 ATTGTTCTGTAGGGGTTTGATTCTTTAAATGAACGAGTGATCAAGAAGAGATATGCTA 394346

QY      226 GATTGCGTAGCGGACGGTGGCGCGCATTCGGCGCTGCTCGCA 269
Db      394345 GATAGGCTTGGGATTGAATGATAGCCGTTTATCTTGCA 394302
```

Search completed: June 19, 2004, 07:23:01
Job time : 46.2924 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 06:08:35 ; Search time 156.115 Seconds
(without alignments)
9683.657 Million cell updates/sec

Title: US-10-658-691-2
Perfect score: 330
Sequence: 1 atgacccgaaaggtccggt.....tgccgaagcacatcatga 330

Scoring table: GAPDP 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	330	100.0	330	17	US-10-658-691-2
2	79.4	24.1	318	15	US-10-330-662-4
3	79.4	24.1	318	15	US-10-330-661-4
4	65.6	19.9	306	9	US-09-738-626-610
5	65.6	19.9	318	9	US-09-738-626-612
6	65.6	19.9	3309400	9	US-09-738-626-612
7	60.4	18.3	250	10	US-09-535-459-815
8	60.4	18.3	458	10	US-09-918-995-25889
9	60.4	18.3	473	10	US-09-918-995-20620
10	60.4	18.3	554	10	US-09-535-459-431
11	60.4	18.3	958	16	US-10-216-464-12
12	60.4	18.3	958	16	US-10-216-464-19
13	58.8	17.8	294	10	US-09-535-459-430
14	58.2	17.6	413	9	US-09-960-352-7903

15	58.2	17.6	443	9	US-09-960-352-14265	Sequence 14265, A
16	57.4	17.4	520	10	US-09-535-459-432	Sequence 432, App
17	57.2	17.3	296	10	US-09-535-459-429	Sequence 429, App
18	56.6	17.2	248	10	US-09-535-459-819	Sequence 819, App
19	45	13.6	297	10	US-09-535-459-427	Sequence 427, App
20	44.4	13.5	261	10	US-09-535-459-96	Sequence 96, Appli
21	44	13.3	612	15	US-10-032-585-6479	Sequence 6479, Ap
22	43.4	13.2	284	10	US-09-535-459-103	Sequence 103, App
23	43.4	13.2	627	15	US-10-156-761-3892	Sequence 3892, Ap
c	43.4	13.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
24	43.4	13.1	1248	15	US-10-156-761-2826	Sequence 2826, Ap
25	43.2	13.1	2205	16	US-10-369-493-34363	Sequence 34363, A
c	43.2	13.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
26	43	13.0	273	10	US-09-535-459-813	Sequence 813, App
27	43	13.0	283	10	US-09-535-459-816	Sequence 816, App
28	43	13.0	301	10	US-09-535-459-822	Sequence 822, App
29	43	13.0	301	10	US-09-535-459-822	Sequence 822, App
30	41.4	12.5	2652	16	US-10-369-493-42250	Sequence 42250, A
31	40.6	12.3	354	11	US-09-864-408A-1511	Sequence 1511, Ap
32	40.6	12.0	1389	15	US-10-156-761-5103	Sequence 5103, Ap
c	39.6	12.0	3807	13	US-10-282-122A-26053	Sequence 26053, A
33	39.6	12.0	3807	13	US-10-282-122A-26053	Sequence 26053, A
34	38.8	11.8	212	10	US-09-535-459-804	Sequence 804, App
35	38.8	11.8	938	13	US-10-425-114-24676	Sequence 24676, A
36	38.8	11.8	1431	13	US-10-425-114-26383	Sequence 26383, A
37	38.6	11.7	888	13	US-10-425-114-23574	Sequence 23574, A
c	38.6	11.7	1522	13	US-10-425-114-33497	Sequence 33497, A
38	38.6	11.7	1522	13	US-10-425-114-33497	Sequence 33497, A
39	38.6	11.7	1534	13	US-10-425-114-31012	Sequence 31012, A
40	38.6	11.7	1534	13	US-10-425-114-31012	Sequence 31012, A
c	38.4	11.6	1830121	15	US-10-329-960-1	Sequence 1, Appli
41	38.4	11.6	1830121	15	US-10-329-960-1	Sequence 1, Appli
c	38.4	11.6	1830121	16	US-10-389-647-332	Sequence 332, App
42	38	11.5	1266	15	US-10-127-032-38	Sequence 38, Appli
c	38	11.5	1266	15	US-10-127-032-38	Sequence 38, Appli

ALIGNMENTS

RESULT 1
US-10-658-691-2
; Sequence 2, Application US/10658691
; Publication No. US20040110258A1
; GENERAL INFORMATION:
; APPLICANT: Kayser, Kevin J.
; APPLICANT: Kilbane, John J.
; TITLE OF INVENTION: Method for Metabolizing Carbazole in Petroleum
; FILE REFERENCE: GTI-1512
; CURRENT APPLICATION NUMBER: US/10/658,691
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: US 60/409,562
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Sphingomonas sp.
US-10-658-691-2

Query Match	100.0%;	Score 330;	DB 17;	Length 330;
Best Local Similarity	100.0%;	Pred. No. 2.2e-91;		
Matches 330;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
Qy	1	ATGACCGCAAGGTCCGCGTGATCTTCGCGAGCGCGCGGCTTCGAGCATCTGTCGAA	60	
Db	1	ATGACCGCAAGGTCCGCGTGATCTTCGCGAGCGCGCGGCTTCGAGCATCTGTCGAA	60	
Qy	61	ACCGAAGCGGAGTATCGCTCATGGAAGCGCGCGCTTCTGAAACCGCGTATCGAA	120	
Db	61	ACCGAAGCGGAGTATCGCTCATGGAAGCGCGCGCTTCTGAAACCGCGTATCGAA	120	
Qy	121	GCGTTTGGGGGGCGCTGTGCTCGCCACGCTTACGTTGCCCCGAGTGG	180	
Db	121	GCGTTTGGGGGGCGCTGTGCTCGCCACGCTTACGTTGCCCCGAGTGG	180	

```
QY 181 CTAGATGCGGTGAACCGCGAGTGAGACCGAAGACGAATGCTCGATTGCGTAGCGGAA 240
Db 181 CTAGATGCGGTGAACCGCGAGTGAGACCGAAGACGAATGCTCGATTGCGTAGCGGAA 240
QY 241 CGTGCCTGATTCGCGGTGCTCTCCAGATCCGCTTACCGAAGCTGCTGACGCGCTG 300
Db 241 CGTGCCTGATTCGCGGTGCTCTCCAGATCCGCTTACCGAAGCTGCTGACGCGCTG 300
QY 301 ACCTGGAATGCGCGAAGGCACATGTCATGA 330
Db 301 ACCTGGAATGCGCGAAGGCACATGTCATGA 330

RESULT 2
US-10-330-662-4
; Sequence 4, Application US/10330662
; Publication No. US20030115626A1
; GENERAL INFORMATION:
; APPLICANT: Weeks, Donald
; APPLICANT: Wang, Xiao-Zhuo
; APPLICANT: Herman, Patricia
; TITLE OF INVENTION: Methods and Materials for Making and Using Transgenic Dicamba-Deg
; TITLE OF INVENTION: Organisms
; FILE REFERENCE: 3553-18-1
; CURRENT APPLICATION NUMBER: US/10/330,662
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/797,238
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/055,145
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/042,666
; PRIOR FILING DATE: 1997-04-04
; PRIOR APPLICATION NUMBER: 60/042,941
; PRIOR FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Pseudomonas maltophilia DI-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(318)
US-10-330-662-4

Query Match 24.1%; Score 79.4; DB 15; Length 318;
Best Local Similarity 55.7%; Pred. No. 1.4e-14;
Matches 152; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 55 GTCGAAACCGAAGCGGAGTATCGCTCATGGAAGCGGCGCTTCTGAAACGGCGTGGACGGT 114
Db 46 GTCGAGCGAGTGAAGCGCGCACCCCTGATGGAAGTCAATCCGACACAGCGGTTTTCAGCAA 105
QY 115 ATCGAAGCGCTTTGCGGGCGCGCTGTGCTGCGCCACGTCGCGACGTTTACGTTGGCCCC 174
Db 106 CTCCTGCGCTTTGCGGGCGCGCTGTGCTGCGCGACTGCGACATCGACCCG 165
QY 175 GAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGAAATGCTCGATTGCGTA 234
Db 166 GCCTTCATGACAACTGCGCGAGATGAGCGAAGATGAGCGAAGACGACCTGCTTCGACAGCTCG 225
QY 235 GCGAAGCTGCGCGCATTCGCGCTGCTGCGGATCCGCTTACCGACCTCGCTTACCGACCTGCTCGAC 294
Db 226 GACCACCGCAACGAGTACTCGCGTCTCTGTCGAGATTCGGTTCACCGCGGCGCTCGAA 285
QY 295 GGCCTGACCTTGGAACTGCGAAGGCACAGTCA 327
Db 286 GGCATCAAGTGACGATCGCGGAGAGACTGA 318

Query Match 24.1%; Score 79.4; DB 15; Length 318;
Best Local Similarity 55.7%; Pred. No. 1.4e-14;
Matches 152; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 55 GTCGAAACCGAAGCGGAGTATCGCTCATGGAAGCGGCGCTTCTGAAACGGCGTGGACGGT 114
Db 46 GTCGAGCGAGTGAAGCGCGCACCCCTGATGGAAGTCAATCCGACACAGCGGTTTTCAGCAA 105
QY 115 ATCGAAGCGCTTTGCGGGCGCGCTGTGCTGCGCCACGTCGCGACGTTTACGTTGGCCCC 174
Db 106 CTCCTGCGCTTTGCGGGCGCGCTGTGCTGCGCGACTGCGACATCGACCCG 165
QY 175 GAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGAAATGCTCGATTGCGTA 234
Db 166 GCCTTCATGACAACTGCGCGAGATGAGCGAAGATGAGCGAAGACGACCTGCTTCGACAGCTCG 225
QY 235 GCGAAGCTGCGCGCATTCGCGCTGCTGCGGATCCGCTTACCGACCTGCTTCGAC 294
Db 226 GACCACCGCAACGAGTACTCGCGTCTCTGTCGAGATTCGGTTCACCGCGGCGCTCGAA 285
QY 295 GGCCTGACCTTGGAACTGCGAAGGCACAGTCA 327
Db 286 GGCATCAAGTGACGATCGCGGAGAGACTGA 318

RESULT 3
US-10-330-661-4
; Sequence 4, Application US/10330661
```

```
; Publication No. US20030135879A1
; GENERAL INFORMATION:
; APPLICANT: Weeks, Donald
; APPLICANT: Wang, Xiao-Zhuo
; APPLICANT: Herman, Patricia
; TITLE OF INVENTION: Methods and Materials for Making and Using Transgenic Dicamba-Deg
; TITLE OF INVENTION: Organisms
; FILE REFERENCE: 3553-18-1
; CURRENT APPLICATION NUMBER: US/10/330,661
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/797,238
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/055,145
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/042,666
; PRIOR FILING DATE: 1997-04-04
; PRIOR APPLICATION NUMBER: 60/042,941
; PRIOR FILING DATE: 1997-04-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Pseudomonas maltophilia DI-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(318)
US-10-330-661-4

Query Match 24.1%; Score 79.4; DB 15; Length 318;
Best Local Similarity 55.7%; Pred. No. 1.4e-14;
Matches 152; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 55 GTCGAAACCGAAGCGGAGTATCGCTCATGGAAGCGGCGCTTCTGAAACGGCGTGGACGGT 114
Db 46 GTCGAGCGAGTGAAGCGCGCACCCCTGATGGAAGTCAATCCGACACAGCGGTTTTCAGCAA 105
QY 115 ATCGAAGCGCTTTGCGGGCGCGCTGTGCTGCGCCACGTCGCGACGTTTACGTTGGCCCC 174
Db 106 CTCCTGCGCTTTGCGGGCGCGCTGTGCTGCGCGACTGCGACATCGACCCG 165
QY 175 GAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGAAATGCTCGATTGCGTA 234
Db 166 GCCTTCATGACAACTGCGCGAGATGAGCGAAGATGAGCGAAGACGACCTGCTTCGACAGCTCG 225
QY 235 GCGAAGCTGCGCGCATTCGCGCTGCTGCGGATCCGCTTACCGACCTCGCTTACCGACCTGCTCGAC 294
Db 226 GACCACCGCAACGAGTACTCGCGTCTCTGTCGAGATTCGGTTCACCGCGGCGCTCGAA 285
QY 295 GGCCTGACCTTGGAACTGCGAAGGCACAGTCA 327
Db 286 GGCATCAAGTGACGATCGCGGAGAGACTGA 318

RESULT 4
US-09-738-626-610/c
; Sequence 610, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
```



```
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 610
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-610

Query Match      19.9%; Score 65.6; DB 9; Length 306;
Best Local Similarity 58.2%; Pred. No. 2.4e-10;
Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

QY 56 TCGAAACCGAGCGGAGTATCGCTCATGGAAGCGCGCGTTCGAAACGGCGTGGACGGTA 115
Db |||||
286 TCGAGGCGACTGTGGTGATTACGTAATGGAGACCGCAGTCCGAAACGGAGTGCCTGGAA 227
QY 116 TCGAAGCCGTTTGGCGGGCGCCTGTGCTGCGCCACGTCGACGTTTACGTTGGCCCG 175
Db |||||
226 TTGTTGCTGAATGCGGCGGTTCCCTTATCGTGTGCAACCTGCCATGTTTGTGACCCG 167
QY 176 AGTGCTAGATGCGCTGAAACCGCGGAGTGAGACCGAAGACGAAATGCT---CGATTGG 232
Db |||||
165 CACAGTATGATGCGCTTCCCCCAATGGAGAGATGGAAGATGAAATGCTGTGGGTGCTG 107
QY 233 TAGCGGACGTCGCGCCGATTGCGGGCTGCTCGCCAGATCCGCTTACCGA 284
Db |||||
106 CCGTGGACCGTGAGGATTGCTCCCGTTTGTCTTCCCAATCAAGGTCAACCGA 55

RESULT 5
US-09-738-626-612
; Sequence 612, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      19.9%; Score 65.6; DB 9; Length 3309400;
Best Local Similarity 58.2%; Pred. No. 5.2e-10;
Matches 135; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

QY 56 TCGAAACCGAGCGGAGTATCGCTCATGGAAGCGCGCGTTCGAAACGGCGTGGACGGTA 115
Db 562917 TCGAGGCGACTTTGCTGATTACGTAATGGAGACCGCAGTCCGAAACGGAGTGCCTGGAA 562858
QY 116 TCGAAGCCGTTTGGCGGGCGCCTGTGCTGCGCCACGTCGACGTTTACGTTGGCCCG 175
Db 562857 TTGTTGCTGAATGCGGCGGTTCCCTTATCGTGTGCAACCTGCCATGTTTGTGACCCG 562798
QY 176 AGTGCTAGATGCGCTGAAACCGCGGAGTGAGACCGAAGACGAAATGCT---CGATTGG 232
Db 562797 CACAGTATGATGCGCTTCCCCCAATGGAGAGATGGAAGATGAAATGCTGTGGGTGCTG 562738
QY 233 TAGCGGAAACGTCGCGCCGATTGCGGGCTGCTCGCCAGATCCGCTTACCGA 284
Db 562737 CCGTGGACCGTGAGGATTGCTCCCGTTTGTCTTCCCAATCAAGGTCAACCGA 562686

RESULT 7
US-09-535-459-815
; Sequence 815, Application US/09535459
```

```
; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
; FILE REFERENCE: PD-1014 CIP
; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 815
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 hu00434840
US-09-535-459-815

Query Match      18.3%; Score 60.4; DB 10; Length 250;
Best Local Similarity 55.1%; Pred. No. 9.8e-09;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 105 CGTGGACGGTATCGAAGCCGTTTTCGGGGCGCGCTGTGCTGCGCCACGTGCCACGTTTA 164
Db 8 CGGGGTGGACCTGGAAGGGCCCTGTGAAGCGCTCCCTGGCCCTGTCTCCACCTGCCATGTGTA 67

QY 165 CGTTGCCCCGAGTGTAGTGGCTGAAACCGCCGAGTGTGATGAGACGGAAGAAATGCT 224
Db 68 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGAGGGAAGACGACATGCT 127

QY 225 CGATTGCGTAGCGAAGCTGCGCGCATTCGGGGCTGCTCCAGATCCGCTTACCGA 284
Db 128 AGACATGCCCCCTCTCCAGAGAACTCGCGCTGGGCTGCCAGATTGCTGACACC 187

QY 285 CTGCTCGACGGCTGACCTGGAATGCCGAAG 318
Db 188 GGAGCTGGAAGGAGCGGAATTCACCCCTGCCCAAG 221

RESULT 8
US-09-918-995-25889
; Sequence 25889, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25889
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-25889

Query Match      18.3%; Score 60.4; DB 10; Length 458;
Best Local Similarity 55.1%; Pred. No. 9.8e-09;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/535,459
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2170
; SOFTWARE: PERL Program
; SEQ ID NO 815
; LENGTH: 250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030040615A1 hu00434840
US-09-535-459-815

Query Match      18.3%; Score 60.4; DB 10; Length 250;
Best Local Similarity 55.1%; Pred. No. 9.8e-09;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 105 CGTGGACGGTATCGAAGCCGTTTTCGGGGCGCGCTGTGCTGCGCCACGTGCCACGTTTA 164
Db 8 CGGGGTGGACCTGGAAGGGCCCTGTGAAGCGCTCCCTGGCCCTGTCTCCACCTGCCATGTGTA 67

QY 165 CGTTGCCCCGAGTGTAGTGGCTGAAACCGCCGAGTGTGATGAGACGGAAGAAATGCT 224
Db 68 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGAGGGAAGACGACATGCT 127

QY 225 CGATTGCGTAGCGAAGCTGCGCGCATTCGGGGCTGCTCCAGATCCGCTTACCGA 284
Db 128 AGACATGCCCCCTCTCCAGAGAACTCGCGCTGGGCTGCCAGATTGCTGACACC 187

QY 285 CTGCTCGACGGCTGACCTGGAATGCCGAAG 318
Db 188 GGAGCTGGAAGGAGCGGAATTCACCCCTGCCCAAG 221

RESULT 9
US-09-918-995-20620
; Sequence 20620, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20620
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(473)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-20620

Query Match      18.3%; Score 60.4; DB 10; Length 473;
Best Local Similarity 55.1%; Pred. No. 9.8e-09;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 105 CGTGGACGGTATCGAAGCCGTTTTCGGGGCGCGCTGTGCTGCGCCACGTGCCACGTTTA 164
Db 151 CGGGGTGGACCTGGAAGGGCCCTGTGAAGCCTCCCTGGCCCTGTCTCCACCTGCCATGTGTA 210

QY 165 CGTTGCCCCGAGTGTAGTGGCTGAAACCGCCGAGTGTGATGAGACGGAAGAAATGCT 224
Db 211 TGTGAGTGAAGACCACTGGATCTCTGCTCTCTCCGAGGAGAGGGAAGACGACATGCT 270

QY 225 CGATTGCGTAGCGAAGCTGCGCGCATTCGGGGCTGCTCCAGATCCGCTTACCGA 284
Db 271 AGACATGCCCCCTCTCCAGAGAACTCGCGCTGGGCTGCCAGATTGCTGACACC 330

QY 285 CTGCTCGACGGCTGACCTGGAATGCCGAAG 318
Db 331 GGAGCTGGAAGGAGCGGAATTCACCCCTGCCCAAG 364

RESULT 10
US-09-535-459-431
; Sequence 431, Application US/09535459
; Publication No. US20030040615A1
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
```

APPLICANT: Mullaby, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSPORT MOLECULE
FILE REFERENCE: PD-1014 CIP
CURRENT APPLICATION NUMBER: US/09/535,459
CURRENT FILING DATE: 2000-03-24
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 2170
SOFTWARE: PERL Program
SEQ ID NO 431
LENGTH: 554
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030040615A1 hu01033738
US-09-535-459-431

Query Match 18.3%; Score 60.4; DB 10; Length 554;
Best Local Similarity 55.1%; Pred. No. 1e-08;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 105 CQTGACGGTATCGAAGCCGTTTGGGGGGGCGCTGTGCTGCGCCACGTCGCGCAAGTCTTA 164
Db 121 CGGGGTGACCTCGAAGGGGCGCTGTGAAGCCTCCCTGGCCTGCTCCACCTGCCATGTGTA 180
QY 165 CTTGGCCCCCGAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCT 224
Db 181 TGTGAGTGAAGACCACTGGATCTCTGCTCTCCGAGGAGAGGGAAGACGATGCT 240
QY 225 CGATTGCGTAGCGGAACCTGCGCCGATTCGCGGCTGCTGCGCAGATCCGCCCTTACCGA 284
Db 241 AGACATGGCCCCCTCCAGGAGAACTCGCGGCTGGGCTGCCAGATTGTGCTGACACC 300
QY 285 CTTGCTGACGCGCTGACCTCGAAGTCCGCGAAG 318
Db 301 GGAGCTGGAAGAGCGGAATTACCTGCGCCAAG 334

RESULT 11
US-10-216-464-12
Sequence 12, Application US/10216464
Publication No. US20030207285A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT219C1
CURRENT APPLICATION NUMBER: US/10/216,464
CURRENT FILING DATE: 2002-08-12
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12

LENGTH: 958
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (873)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (916)
OTHER INFORMATION: n equals a,t,g, or c
US-10-216-464-12
Query Match 18.3%; Score 60.4; DB 16; Length 958;
Best Local Similarity 55.1%; Pred. No. 1e-08;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 105 CQTGACGGTATCGAAGCCGTTTGGGGGGGCGCTGTGCTGCGCCACGTCGCGCAAGTCTTA 164
Db 291 CGGGGTGACCTCGAAGGGGCGCTGTGAAGCCTCCCTGGCCTGCTCCACCTGCCATGTGTA 350
QY 165 CTTGGCCCCCGAGTGGCTAGATGCGCTGAAACCGCGAGTGAGACCGAAGACGAAATGCT 224
Db 351 TGTGAGTGAAGACCACTGGATCTCTGCTCTCCGAGGAGAGGGAAGACGATGCT 410
QY 225 CGATTGCGTAGCGGAACGTCGCGCGCAATTCGCGGCTGCTGCGCAGATCCGCCCTTACCGA 284
Db 411 AGACATGGCCCCCTCCAGGAGAACTCGCGGCTGGGCTGCCAGATTGTGCTGACACC 470
QY 285 CTTGCTGACGCGCTGACCTCGAAGTCCGCGAAG 318
Db 471 GGAGCTGGAAGAGCGGAATTACCTGCGCCAAG 504

RESULT 12
US-10-216-464-19
Sequence 19, Application US/10216464
Publication No. US20030207285A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT219C1
CURRENT APPLICATION NUMBER: US/10/216,464
CURRENT FILING DATE: 2002-08-12
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 958
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature

LOCATION: (873)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (916)
OTHER INFORMATION: n equals a,t,g, or c
US-10-216-464-19

Query Match 18.3%; Score 60.4; DB 16; Length 958;
Best Local Similarity 55.1%; Pred. No. 1e-08;
Matches 118; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 105 CGTGGAGGATCGAAGCCGTTTGGGGGGCGCTGTGCTGCGCCACGTCGACGTTTA 164
DB 291 CGGGTGGAGCTGGAAGGGCCCTGTGAAGCTCCCTGGCTGTCTCCACCTGCATGTGA 350
QY 165 CGTTGGCCCCGAGTGGCTAGATCGCTGGAACCGCCGAGTGAGACCGAAGAAATGCT 224
DB 351 TGTGAGTGAAGACCACTGGATCTCTGCTCTCCGAGGAGGGAAGACACATGCT 410
QY 225 CGATTGGTAGCGAAGCTGCGCGGCAATTCGGGCTGTCTGCCAGATCGGCTTACCGA 284
DB 411 AGACATGCCCCCTCTCCAGAGAACTCGCGCTGGGCTGCCAGATTGTGCTGACACC 470
QY 285 CTTGCTCGAGCGGCTGACCCCTGGAATTCGCCGAAG 318
DB 471 GGAGCTGGAAGGAGCGGAATTCACCTTGCCCAAG 504

RESULT 13
US-09-535-459-430
Sequence 430, Application US/09535459
Publication No. US20030040615A1
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Deleage, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING ELECTRON TRANSFER MOLECULE
FILE REFERENCE: PD-1014 CIP
CURRENT APPLICATION NUMBER: US/09/535,459
Prior application data removed - consult PALM or file wrapper
CURRENT FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 2170
SOFTWARE: PERL Program
SEQ ID NO 430
LENGTH: 294
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030040615A1 hu01270499
US-09-535-459-430

Query Match 17.8%; Score 58.8; DB 10; Length 294;
Best Local Similarity 54.7%; Pred. No. 2.9e-08;
Matches 117; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 105 CGTGGAGGATCGAAGCCGTTTGGGGGGCGCTGTGCTGCGCCACGTCGACGTTTA 164
DB 37 CGGGTGGACCTGGAAGGGCCCTGTGAAGCTCCCTGGCTGTCTCCACCTGCATGTGA 96
QY 165 CGTTGGCCCCGAGTGGCTAGATGGCTGGAACCGCCGAGTGAGACCGAAGAAATGCT 224
DB 97 TGTGAGTGAAGACCACTGGATCTCTTACCTCTCCGAGGAGGGAAGACACATGCT 156
QY 225 CGATTGGTAGCGAAGCTGCGCGCAATTCGGGCTGTCTGCCAGATCGGCTTACCGA 284
DB 157 AGACATGCCCCCTCTCCAGAGAACTCGCGCTGGGCTGCCAGATTGTGCTGACACC 216
QY 285 CTTGCTCGAGCGGCTGACCCCTGGAATTCGCCGAAG 318

DB 217 GGAGCTGGAAGGAGCGGAATTCACCTGCCCAAG 250

RESULT 14
US-09-960-352-7903
Sequence 7903, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Mengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 7903
LENGTH: 413
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 34-LIB3057-016-Q1-K1-A6
US-09-960-352-7903

Query Match 17.6%; Score 58.2; DB 9; Length 413;
Best Local Similarity 55.1%; Pred. No. 4.6e-08;
Matches 114; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 112 GGTATCGAAGCCGTTTGGGGGGCGCTGTGCTGCGCCACGTCGACGTTTACGTTGGC 171
DB 131 GATCTGGAAGGGGCTCGAAGCGTCTTGGCGTCTCCACCTGCCACGTTATGTGAGT 190
QY 172 CCCGAGTGGCTAGATGGCTGAAACCGCCGAGTGAGACCGAAGAAATGCTCGATTGC 231
DB 191 GAGGACCACTGGACCTTCTGCCGCTCTGATGAGAGGGAGGACGACATGCTGGATG 250
QY 232 GTAGCGAAGCTGCGCGCAATTCGCGGCTGTCTTCCAGATCCGCTTACCGACCTGCTC 291
DB 251 GCCCTCTCTCCAGAGAACTCCCGCTGGCTGCCAGATCGTGTGAGCGCTGAGCTG 310
QY 292 GACGCGCTGACCTGGAACTGCCGAAG 318
DB 311 GAAGGGCCGAATTCACCTGCCCAAG 337

RESULT 15
US-09-960-352-14265
Sequence 14265, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Mengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 14265
LENGTH: 443
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 61-LIB3057-009-Q1-K1-H2
US-09-960-352-14265

Query Match 17.6%; Score 58.2; DB 9; Length 443;
Best Local Similarity 55.1%; Pred. No. 4.6e-08;
Matches 114; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 112 GGTATCGAAGCCGTTTGGGGGGCGCCCTGTGCTCGGCCACGTCGCCACGTTTACGTTGGC 171
Db |||||
QY 130 GATCTGGAAGGGCCCTGGGAAGCGTCCTTGGCGTCTCCACCTGCCACGTTGATGTGAGT 189
Db |||||
QY 172 CCCGAGTGGCTAGATGCGCTGAACCCGCGAGTGAGACCGAAGACGAATGCTCGATTGC 231
Db |||||
QY 190 GAGGACCACTGGACCTTCTGCGGCTCTGTATGAGGGAGGACGACATGCTGGATATG 249
Db |||||
QY 232 GTAGCGGAACGTGCGCGCCATTGCGGGCTGTCTTCCAGATCCGCTTACCGACCTGCTC 291
Db |||||
QY 250 GCCCTCTCTCCAAAGAGAACTCCCGGCTGGGCTGCCAGATCGTGTGCTGACGCTGAGCTG 309
Db |||||
QY 292 GACGCGCTGACCCCTGGAACCTGCGAAG 318
Db |||||
QY 310 GAAAGGGCGGAATTCACCTGCCCAAG 336
Db |||||

Search completed: June 19, 2004, 09:16:55
Job time : 171.115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 04:04:20 ; Search time 4587.53 Seconds
(without alignments)
10742.394 Million cell updates/sec

Title: US-10-658-691-1

Perfect score: 1137

Sequence: 1 gtggtaaccaacatcaat.....gcattcaggagcaggtga 1137

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hgt_hum.*

31: em_hgt_inv.*

32: em_hgt_other.*

33: em_hgt_mus.*

34: em_hgt_pln.*

35: em_hgt_rod.*

36: em_hgt_mam.*

37: em_hgt_vrt.*

38: em_sy.*

39: em_hgt_hum.*

40: em_hgt_mus.*

41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1137	100.0	4154	1	AF442494	AF442494 Shingomo
2	1137	100.0	13651	1	AB095953	AB095953 Shingomo
3	859.2	75.6	1110	1	AB072827	AB072827 Shingomo
4	422.6	37.2	20017	1	AB095952	AB095952 Janthinob
5	419.4	36.9	1161	6	E17292	E17292 gDNA encodi
6	419.4	36.9	1200	6	E17293	E17293 gDNA encodi
7	419.4	36.9	6880	1	AB001723	AB001723 Pseudomon
8	419.4	36.9	6881	1	D89064	D89064 Pseudomonas
9	419.4	36.9	44266	1	AB047548	AB047548 Pseudomon
10	419.4	36.9	199035	1	AB088420	AB088420 Pseudomon
11	281.8	24.8	8096	1	AB094786	AB094786 Pseudomon
12	153.6	13.5	3203	1	PPY12655	Y12655 P. putida ox
13	49	4.3	2000	6	AX655393	AX655393 Sequence
14	48.8	4.3	298450	1	SC039107	AL933107 Streptomy
15	47.4	4.2	125020	9	AF429315	AF429315 Homo sapi
16	47	4.1	125020	9	AF429315	AF429315 Homo sapi
17	46.2	4.1	6564	6	AX069073	AX069073 Sequence
18	46.2	4.1	32870	1	AE007122	AE007122 Mycobacte
19	46.2	4.1	318050	1	EX248344	EX248344 Mycobacte
20	46.2	4.1	348676	15	BX842581	BX842581 Mycobacte
21	44	3.9	5727	1	REU24277	U24277 Rhodococcus
22	44	3.9	210205	1	AY23810	AY23810 Rhodococc
23	43	3.8	344321	1	EX640429	EX640429 Bordetell
24	43	3.8	348624	1	EX640441	EX640441 Bordetell
25	42.8	3.8	2000	6	AX655393	AX655393 Sequence
26	41.6	3.7	5957	1	RERBP41	D32142 Rhodococcus
27	41	3.6	15982	1	AE004779	AE004779 Pseudomon
28	41	3.6	89153	1	NSP561198	AX7561198 Actinomad
29	40.6	3.6	15663	1	AE007046	AE007046 Mycobacte
30	40.6	3.6	291050	1	EX248340	EX248340 Mycobacte
31	40.6	3.6	301617	1	AE016911	AE016911 Chromobac
32	40.6	3.6	347496	1	EX842577	EX842577 Mycobacte
33	40.6	3.6	348171	1	EX640412	EX640412 Bordetell
34	40.2	3.5	14886	1	TAU52917	U52917 Thermus aqu
35	40.2	3.5	349640	1	EX572600	EX572600 Rhodopsu
36	40	3.5	3151	1	AB055506	AB055506 Streptomy
37	40	3.5	4483	1	PPPOB4B	X78823 P. pseudoalc
38	40	3.5	6835	1	AF452376	AF452376 Rhodococc
39	40	3.5	20021	1	AE004730	AE004730 Pseudomon
40	40	3.5	299350	1	AP006572	AP006572 Gloeobact
41	39.8	3.5	606	6	AX196058	AX196058 Sequence
42	39.8	3.5	109519	6	AX195929	AX195929 Sequence
43	39.4	3.5	10177	1	AE004977	AE004977 Halobacte
44	39.4	3.5	304500	1	AP005953	AP005953 Bradyrhiz
45	39.2	3.4	15873	1	AB007115	AB007115 Mycobacte

ALIGNMENTS

RESULT 1
AF442494
LOCUS
DEFINITION Shingomonas sp. GTIN11 carbazole operon, partial sequence.
ACCESSION AF442494
VERSION AF442494.1 GI:17227030
KEYWORDS
SOURCE Shingomonas sp. GTIN11
ORGANISM Shingomonas sp. GTIN11
Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
Sphingomonadaceae; Sphingomonas.
REFERENCE
1 (bases 1 to 4154)
AUTHORS
TITLE
Isolation and characterization of Shingomonas sp. GTIN11 capable
of carbazole metabolism in petroleum

JOURNAL	Biochem. Biophys. Res. Commun. 297 (2), 242-248 (2002)
MEDLINE	22222130
PUBMED	12237109
REFERENCE	2 (bases 1 to 4154)
AUTHORS	Kayser, K.J., Daram, A., Abassian, J. and Kilbane, J.J. III.
TITLE	Direct Submission
JOURNAL	Submitted (01-NOV-2001) Biotechnology, Gas Technology Institute, 1700 South Mount Prospect Road, Des Plaines, IL 60018, USA
FEATURES	Location/Qualifiers
source	1. .4154 /organism="Sphingomonas sp. GTIN11" /mol_type="genomic DNA" /db_xref="taxon:174635"
gene	<1. .>4154 /gene="carbazole operon"
gene	650. 1786 /gene="carAa"
CDS	650. .1786 /gene="carAa" /codon_start=1 /transl_table=11 /product="carbazole dioxygenase" /protein_id="AAL37976.1" /db_xref="GI:17227031"
translation	/translations="MANQSIARERTKVNEPIRAKLGFRNHWYFVRLASEIAGETPV FVKGELKILNVRGKVAIQDCLRHGVTLSDRVCEYKNTISCWYHGWTYRWDG RLVDILNTPGSVQIGRRALKTFVEEAKGLIFVYDGCGETPLIEDVPPGLDENRAI HGQRLVAVNWRGLAGNDFGADHFIRKNSILVKGNDIILPLGFAPGPDQDLTRSEVA AGPKGVYDLGSHSVFVPEGMIEGKPAIHGNIGSRVAISISIMLPGLVKVPEWPD ELTQFEWYVVDTSLSLYFTLCKVVTSEKADSPEFEKHWGLALINGFNDDDDIMA RESMEPPYADDRGWSBEILFEPPDRAIIEWRGLASQHNRCIQEAR"
gene	1735. .2067 /gene="carBa"
CDS	1735. .2067 /gene="carBa" /notes="meta cleavage enzyme subunit" /codon_start=1 /transl_table=11 /product="2-aminobiphenyl-2,3-diol 1,2-dioxygenase" /protein_id="AAL37977.1" /db_xref="GI:17227032"
translation	/translations="MAGACQSAQSRHSGSLTMGTASYPVNRIMQELFTNPGNVELFR ADREALYERGLSSAQRALDEGGFALTAUGLHPVLQMHFMLNPNMADPVSVKAY RKNVDRNG"
gene	2060. .2863 /gene="carBb"
CDS	2060. .2863 /gene="carBb" /notes="meta cleavage enzyme catalytic subunit" /codon_start=1 /transl_table=11 /product="2-aminobiphenyl-2,3-diol 1,2-dioxygenase" /protein_id="AAL37978.1"
translation	/translations="MAEIVGFATSHVLFSPDGAQALRVVEGNEIGRRVRLRPD LLVIGSDHFNITRLQPPFTVGSJDTFTPLGMDIEORPAGNRAFAESICARAAD RFDLAQEELRPHGVNVPVPMFDPDGRIPVPIVYNNATPPPTAARALQGLDIVAE AVGLDLPSHLRVVVATGSLSHWINIPGHGVNAEFDRRVIAAFOSGDMRMLRAIDTE TLKKNAGSGLEIVNWMAAATLPGRAEKIYYEPQWNTGNGGAIIV"
gene	2906. .3730 /gene="carC"
CDS	2906. .3730 /gene="carC" /codon_start=1 /transl_table=11 /product="meta cleavage compound hydrolase" /protein_id="AAL37979.1" /db_xref="GI:17227034"
translation	/translations="MVLAAVAGRSVTVRGNETYHHEGSGDVVVLVHGGAGADSMGN WRGMPVLADRYVIAVDMLGFGRTAKPADPPFVSQAARTDHLAGLGSNVALV GNSNGSAGALVAVRGLVKLMSAGLSKIDPALEPLVGYDFTREGMIRLVRA LITDNOFIDAMIDYRALVSDVPETRAYSATMOWIRDQGLVYEDDYIRRTAPTLLI VNGKLDKVVPLANVAYKELELIGPSWGYIMPDCGHWAMIEHPVDFAITTAFFIAAQ"
ORIGIN	Query Match 100.0%; Score 1137; DB 1; Length 4154; Best Local Similarity 100.0%; Pred. No. 7.7e-250; Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GTGGCTAACCAACCATCAATCGCGAGCGAGACCAAGGTTTGGGAGCCTTATATCCGT 60
DB	650 GTGGCTAACCAACCATCAATCGCGAGCGAGACCAAGGTTTGGGAGCCTTATATCCGT 709
QY	61 GCGAACTCGGGTTCGGAACCATTTGGTATCCCGTTCGCTCGCGAGCGAAATCGCCGAA 120
DB	710 GCGAACTCGGGTTCGGAACCATTTGGTATCCCGTTCGCTCGCGAGCGAAATCGCCGAA 769
QY	121 GGTACTCCGTTCCGTCACAACTCCTCGGAGAGAAATCTGTCAATCGCGTGGCGGC 180
DB	770 GGTACTCCGTTCCGTCACAACTCCTCGGAGAGAAATCTGTCAATCGCGTGGCGGC 829
QY	181 AAGTCTATGATGATCCAGGACAGTGTCTGATCGCGTGTAAAGCTTTCGACCGGGTC 240
DB	830 AAGTCTATGATGATCCAGGACAGTGTCTGATCGCGTGTAAAGCTTTCGACCGGGTC 889
QY	241 GAGTCTATTCAGAAACACCATATCTGTGTGTATCACGGTGGACATATCGTGGGAC 300
DB	890 GAGTCTATTCAGAAACACCATATCTGTGTGTATCACGGTGGACATATCGTGGGAC 949
QY	301 GATGCGCCCTCGTGCATATCTCACAACCCCGGAGTGTGCAGATCGCGCGCGCGCT 360
DB	950 GATGCGCCCTCGTGCATATCTCACAACCCCGGAGTGTGCAGATCGCGCGCGCGCT 1009
QY	361 TTGAAGAGCTTCCCGTTGAAGAGGCCAAAGTCTTATCTTCTGTTTACGTAGGCGAGCG 420
DB	1010 TTGAAGAGCTTCCCGTTGAAGAGGCCAAAGTCTTATCTTCTGTTTACGTAGGCGAGCG 1069
QY	421 GAACCAAGCGCTTATCGAAGATGTGCGCCCGCGCTTCTTGTATGAAGACCGCGCAT 480
DB	1070 GAACCAAGCGCTTATCGAAGATGTGCGCCCGCGCTTCTTGTATGAAGACCGCGCAT 1129
QY	481 CACGCCCAACATCGCTCGTGGCTCGAACTGGCGCTTGGTGCAGAAACCGCTTTGAT 540
DB	1130 CACGCCCAACATCGCTCGTGGCTCGAACTGGCGCTTGGTGCAGAAACCGCTTTGAT 1189
QY	541 GCGGGCACGCTTTCATTCACAAGATTCGATCTGTGAAGGCGACGATATCATCTG 600
DB	1190 GCGGGCACGCTTTCATTCACAAGATTCGATCTGTGAAGGCGACGATATCATCTG 1249
QY	601 CGCTTGGCTTTCGCGCTCGCGATCCGACCGCTTACCGCTTCGAGGTTGCTCGCGGC 660
DB	1250 CGCTTGGCTTTCGCGCTCGCGATCCGACCGCTTACCGCTTCGAGGTTGCTCGCGGC 1309
QY	661 AAGCCCAAGGTTTACGATCTGTTGGCGATTCGCTGCGCGTTTCGAAAGGCATG 720
DB	1310 AAGCCCAAGGTTTACGATCTGTTGGCGATTCGCTGCGCGTTTCGAAAGGCATG 1369
QY	721 ATCGAAGCAACCTGCAATCCATGGCAACATTGGCAACATTGGCAGACCGCTCGCCAT 780
DB	1370 ATCGAAGCAACCTGCAATCCATGGCAACATTGGCAGACCGCTCGCCATCAGCATA 1429
QY	781 TCGATCTGGCTGCCGGCGTACTCAAGTGCAGACCGTGGCGCGATCCCGAGCTCACGAG 840
DB	1430 TCGATCTGGCTGCCGGCGTACTCAAGTGCAGACCGTGGCGCGATCCCGAGCTCACGAG 1489


```
QY      841  TTCGAATGTTAGTCCCGTTCGATGAGACCAACCTCTACTTCCAGACGCTGGGAAA  900
Db      1490  TTCGAATGTTAGTCCCGTTCGATGAGACCAACCTCTACTTCCAGACGCTGGGAAA  1549
QY      901  GTCTGAGCTCAAGAGGAGCGGACGACCTCTTCGAGCGAGAAATCCAGAAAATGGGTA  960
Db      1550  GTCTGAGCTCAAGAGGAGCGGACGACCTCTTCGAGCGAGAAATCCAGAAAATGGGTA  1609
QY      961  GGCCTCGCGCTTAACGGCTTCAATGATGACGACATCATGGACGCGATCATGATGAGCGG  1020
Db      1610  GGCCTCGCGCTTAAACGGCTTCAATGATGACGACATCATGGACGCGATCATGATGAGCGG  1669
QY      1021  TTCTACGCTGATGATCGCGGTTGGTCCGAGAAATCTCTTTCGAGCGGACCGCGCAATC  1080
Db      1670  TTCTACGCTGATGATCGCGGTTGGTCCGAGAAATCTCTTTCGAGCGGACCGCGCAATC  1729
QY      1081  ATCGAGTGGCGGGGCTTCCAGTCAGACATCGCGGATTCAGAGGACGCTTGA  1137
Db      1730  ATCGAGTGGCGGGGCTTCCAGTCAGACATCGCGGATTCAGAGGACGCTTGA  1786

RESULT 2
LOCUS   AB095953                      13651 bp    DNA    linear    BCT 04-FEB-2003
DEFINITION   Sphingomonas sp. KAI car operons, complete cds.
ACCESSION   AB095953
VERSION     AB095953.1  GI:28201219
KEYWORDS
SOURCE      Sphingomonas sp. KAI
ORGANISM    Sphingomonas sp. KAI
            Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
            Sphingomonadaceae; Sphingomonas.
REFERENCE   1
AUTHORS    Inoue,K., Nojiri,H., Nakai,S., Endoh,T., Urata,M., Ashikawa,Y.,
            Saiki,Y., Yoshida,T., Habe,H. and Omori,T.
TITLE      Divergent structures of carbazole-degrading car operons isolated
            from several gram-negative bacteria
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 13651)
AUTHORS    Nojiri,H. and Inoue,K.
TITLE      Direct Submission
JOURNAL    Submitted (11-NOV-2002) Hideaki Nojiri, Biotechnology Research
            Center, The University of Tokyo; 1-1-1, Yayoi, Bunkyo-ku, Tokyo
            113-8657, Japan (E-mail:anojiri@mail.ecc.u-tokyo.ac.jp,
            Tel.81-3-5841-3067, Fax:81-3-5841-8030)
FEATURES   Location/Qualifiers
            1..13651
             /organism="Sphingomonas sp. KAI"
             /mol_type="genomic DNA"
             /strain="KAI"
             /db_xref="taxon:164608"
             complement(127..2199)
             /gene="ORF10"
             complement(127..2199)
             /gene="ORF10"
             /function="unknown"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein"
             /protein_id="BAC56754.1"
             /db_xref="GI:28201220"
             /translation="MISEKIGQHLERKALYVROSSVHQVLHNRESSTLOYAMGRLL
             TANGWSHIETDIDLGSAAGVTRAGDFRWVAEVLGKGVAAAREVSRFARNRSDW
             QOLIEKRVVDVTLDOETVYAPROCDNRLLGLKSLNEVELDLRLQRLSARYEKA
             RQELVVAAPGVFKAGDRIEKDPDRRQVEALVFDKVAELGSAQALLWFIHGLD
             LPAKNGGVWRRPSYATIHMIENPIYGGAAYGKSRATGYDGTAMRSRERKAR
             DWLALLPGSHGYSWSESDIRKNSVDNVTSSHGAPXHGDLALLGLRCRCGR
             KUTRYAGHDIPRYSCGGLDNGEPCFTAGGLRVDDDAERALLQVLEFGCAAS
             VSAEAQADRQDVRLNRDLAARYAADRAFQYDAADPQNRVLAARELETRNRAL
             TEAGEYEARIVAHDASTAPALPSLDIDGLASLEAVNAPQSDARLKKIVRTLQ
             EVVADIDHDASIVILLIHWVGVDLRLPRRRKQGRNSTADIIAAREVLVLIAND
             LIAGILNRNLGTVGHGNNRTRERVTLRSRHHRI PVFRTVADGPAFWLNLQAARHIGV
```

gene	6328..6609	/gene="carBa"	Query Match	100.0%;	Score 1137;	DB 1;	Length 13651;
CDS	6328..6609	/gene="carBa"	Best Local Similarity	100.0%;	Pred. No. 7.6e-250;		
	/codon_start=1	/codon_start=1	Matches 1137;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
	/transl_table=11	/transl_table=11					
	/product="subunit of meta cleavage enzyme"	/product="subunit of meta cleavage enzyme"					
	/protein_id="BAC56760.1"	/protein_id="BAC56760.1"					
	/db_xref="GI:28201226"	/db_xref="GI:28201226"					
	/translation="MGTTASYPVNRLMQELFTNPGNVELFPADREALYERYGLSSAQR	/translation="MGTTASYPVNRLMQELFTNPGNVELFPADREALYERYGLSSAQR					
gene	6602..7405	/gene="carBb"	QY	1	GTGGCTAAACCAACCATCAATCGCGAGCGCAGAAACCAAGGTTTGGAGCCTTATATCCGT	60	
CDS	6602..7405	/gene="carBb"	DB	5192	GTGGCTAACCAACCATCAATCGCGAGCGCAGAAACCAAGGTTTGGAGCCTTATATCCGT	5251	
	/codon_start=1	/codon_start=1					
	/transl_table=11	/transl_table=11					
	/product="subunit of meta cleavage enzyme"	/product="subunit of meta cleavage enzyme"					
	/protein_id="BAC56761.1"	/protein_id="BAC56761.1"					
	/db_xref="GI:28201227"	/db_xref="GI:28201227"					
	/translation="MAEIVGVFATSHVLFSGPDGAQALRVVEGNEIGRVRLLRPD	/translation="MAEIVGVFATSHVLFSGPDGAQALRVVEGNEIGRVRLLRPD					
	LLIVGSHLENIITRLQPTTGVSDTFTPLGMDIEQRFAGNAPAFSLCARAD	LLIVGSHLENIITRLQPTTGVSDTFTPLGMDIEQRFAGNAPAFSLCARAD					
	RFDLAQGELEPDRHGVPLMFIIDPDRIPVPIYVNNMTPTTAAARAAQLGDI	RFDLAQGELEPDRHGVPLMFIIDPDRIPVPIYVNNMTPTTAAARAAQLGDI					
	AVGLDPLSHLRVAVATGCLSHWINIPGHGVNAEFORRVIAAFQSGDMRWLR	AVGLDPLSHLRVAVATGCLSHWINIPGHGVNAEFORRVIAAFQSGDMRWLR					
	LAIDTE	LAIDTE					
	TLLKVAAGGGLVIVNMAAATLPGRAEKIYBPMQWMTGMGIAIV	TLLKVAAGGGLVIVNMAAATLPGRAEKIYBPMQWMTGMGIAIV					
gene	7448..8272	/gene="carC"	QY	121	GGTACTCCCGTTCCTCCGTCAGCTCCTGGGAGAGAGATTCCTCAATCGCTGGCGGC	180	
CDS	7448..8272	/gene="carC"	DB	5312	GGTACTCCCGTTCCTCCGTCAGCTCCTGGGAGAGAGATTCCTCAATCGCTGGCGGC	5371	
	/codon_start=1	/codon_start=1					
	/transl_table=11	/transl_table=11					
	/product="meta cleavage compound hydrolase"	/product="meta cleavage compound hydrolase"					
	/protein_id="BAC56762.1"	/protein_id="BAC56762.1"					
	/db_xref="GI:28201228"	/db_xref="GI:28201228"					
	/translation="WVLAATAGRSVTVRGMEITYHEOGSDVVVHVHGGAGDSXGN	/translation="WVLAATAGRSVTVRGMEITYHEOGSDVVVHVHGGAGDSXGN					
	WRGVPVLAADRVATVAMELGFTAKPADPFFVSCAARTDHLGFLDALGSLNALV	WRGVPVLAADRVATVAMELGFTAKPADPFFVSCAARTDHLGFLDALGSLNALV					
	GNSMGAGALVAVERPGLVRLVMSAGLVSKIDPALEPLVGYDTFREGMIELVRA	GNSMGAGALVAVERPGLVRLVMSAGLVSKIDPALEPLVGYDTFREGMIELVRA					
	LTTDFQDDAMIDYRALSVDPTERRAYSATMQRDOGLLYEDDYIRITAPTII	LTTDFQDDAMIDYRALSVDPTERRAYSATMQRDOGLLYEDDYIRITAPTII					
	VNGKLDKVPVLANAYKFLLELIGPSMGVIMPCDGHAMIEHPVDPTAAPIEAQ	VNGKLDKVPVLANAYKFLLELIGPSMGVIMPCDGHAMIEHPVDPTAAPIEAQ					
gene	8313..8642	/gene="carAc"	QY	421	GAACCAACGCGCTTATCGAAGATGTGCGCGCGCTTCTTGTATGATAAACCGCGCAIT	480	
CDS	8313..8642	/gene="carAc"	DB	5612	GAACCAACGCGCTTATCGAAGATGTGCGCGCGCTTCTTGTATGATAAACCGCGCAIT	5671	
	/codon_start=1	/codon_start=1					
	/transl_table=11	/transl_table=11					
	/product="ferredoxin component of CAR 1,9a-dioxygenase"	/product="ferredoxin component of CAR 1,9a-dioxygenase"					
	/protein_id="BAC56763.1"	/protein_id="BAC56763.1"					
	/db_xref="GI:28201229"	/db_xref="GI:28201229"					
	/translation="WTKRVVIFRAAGGFHELVETEAGVSLMEAAVLNGVDGIEAVCG	/translation="WTKRVVIFRAAGGFHELVETEAGVSLMEAAVLNGVDGIEAVCG					
	GACACATRVYVGPBWDALPFPSETEDMLDCVAERAPHSLRSLCQINLTDLLGLTL	GACACATRVYVGPBWDALPFPSETEDMLDCVAERAPHSLRSLCQINLTDLLGLTL					
	ELPRAQS"	ELPRAQS"					
gene	8684..10999	/gene="ORF11"	QY	601	CCGCTTGGCTTTCGCGCTGGCGATCCCGACAGCTTACGGGTTCCGAGTTGCTCGGGC	660	
CDS	8684..10999	/gene="ORF11"	DB	5792	CCGCTTGGCTTTCGCGCTGGCGATCCCGACAGCTTACGGGTTTCGAGGCATG	5911	
	/function="unknown"	/function="unknown"					
	/codon_start=1	/codon_start=1					
	/transl_table=11	/transl_table=11					
	/product="hypothetical protein"	/product="hypothetical protein"					
	/protein_id="BAC56764.1"	/protein_id="BAC56764.1"					
	/db_xref="GI:28201230"	/db_xref="GI:28201230"					
	/translation="WCKLMSGALVLAASAPAPAOOTDGVEDASLEIITVARR	/translation="WCKLMSGALVLAASAPAPAOOTDGVEDASLEIITVARR					
	VSENLQDSFVALSARGENAIERRAINVSDIGKTPSLSPDSAAAGSSNSVTVLR	VSENLQDSFVALSARGENAIERRAINVSDIGKTPSLSPDSAAAGSSNSVTVLR					
	GTGQDFNMVDPGVGILYDGYISRSVGALLETRDVSQVLRGPQTLFGKNTIG	GTGQDFNMVDPGVGILYDGYISRSVGALLETRDVSQVLRGPQTLFGKNTIG					
	AVITSRPDEQNGNIEATGSDYINRINRGARINPVSEDFRYVSGALTRDGYFR	AVITSRPDEQNGNIEATGSDYINRINRGARINPVSEDFRYVSGALTRDGYFR					
	LVDGQKNGNSYSGRFVAWMDPFDNLLISADYTRERAVATLLDNNENAIPL	LVDGQKNGNSYSGRFVAWMDPFDNLLISADYTRERAVATLLDNNENAIPL					
	FNVPFLNAPGLPPTSPSCYNQMLTGDKNTNPNRVSVDVWVGSLNSWDA	FNVPFLNAPGLPPTSPSCYNQMLTGDKNTNPNRVSVDVWVGSLNSWDA					
	GPVITKSTAYRELSAFVADIDQSPALAIQRTENDYQNSQSFQLOGDGEKLY	GPVITKSTAYRELSAFVADIDQSPALAIQRTENDYQNSQSFQLOGDGEKLY					
	TGLVYHGGGTDINSLFTAPTDFSSGVKNDYVAAFAQFIHATDNLNLGRTY	TGLVYHGGGTDINSLFTAPTDFSSGVKNDYVAAFAQFIHATDNLNLGRTY					
	SRDKFSQDOELVHAPAAQPLPSPFTBPLTNPVTSMAAGSPQLRVDIYVPAV	SRDKFSQDOELVHAPAAQPLPSPFTBPLTNPVTSMAAGSPQLRVDIYVPAV					
	QRTTSANFTPBTLDYKFTPDMLGYFTFSGFKSGGRTQVFPFEPAPAFEPFAR	QRTTSANFTPBTLDYKFTPDMLGYFTFSGFKSGGRTQVFPFEPAPAFEPFAR					
	VLEIGVVELFNRLKSLADFTIKYISDLIIVNDGPAPKRNAGKIRPFESAEQL	VLEIGVVELFNRLKSLADFTIKYISDLIIVNDGPAPKRNAGKIRPFESAEQL					
	VVTDAFRUTGAVSYLDAEYREVDAAPVTVDSALANTPKWTANAGITVEPYDEAGR	VVTDAFRUTGAVSYLDAEYREVDAAPVTVDSALANTPKWTANAGITVEPYDEAGR					

```

QY 961 GGCCTCGCGCTTAAACGGCTTCAATGATGACGACATCATGGACGATGATCGATGAGCGG 1020
Db 6152 GGCCTCGCGCTTAAACGGCTTCAATGATGACGACATCATGGACGATGATCGATGAGCGG 6211
QY 1021 TTCTACGCTGATGATCGCGGTTGGTTCGAGAAATCCTGTTGAGCCGCGGACCGCAATC 1080
Db 6212 TTCTACGCTGATGATCGCGGTTGGTTCGAGAAATCCTGTTGAGCCGCGGACCGCAATC 6271
QY 1081 ATCGAGTGGCGGGGCTTGCAGTGCAGACATCGCGGATTCAGGAAGCAGCTTGA 1137
Db 6272 ATCGAGTGGCGGGGCTTGCAGTGCAGACATCGCGGATTCAGGAAGCAGCTTGA 6328

RESULT 3
AB072827
LOCUS
DEFINITION
  Spingomonas sp. KAI carAa, carBa genes for putative oxygenase
  component of carbazole 1,9a-dioxygenase, putative small subunit of
  2'-aminobiphenyl 2,3-diol, partial cds.
ACCESSION
  AB072827.1 GI:20135960
KEYWORDS
  Spingomonas sp. KAI
SOURCE
  Spingomonas sp. KAI
  Bacteria; Proteobacteria; Alphaproteobacteria; Spingomonadales;
  Spingomonadaceae; Spingomonas.
REFERENCE
  1 Haba,H., Ashikawa,Y., Saiki,Y., Yoshida,T., Nojiri,H. and Omori,T.
    Spingomonas sp. strain KAI, carrying a carbazole dioxygenase gene
    homologue, degrades chlorinated dibenzo-p-dioxins in soil
    FEMS Microbiol. Lett. 211 (1), 43-49 (2002)
  2 22047665
  12052549
  2 (bases 1 to 1110)
  Omori,T., Haba,H. and Nojiri,H.
  Direct Submission
  Submitted (10-OCT-2001) Hiroshi Haba, The University of Tokyo,
  Biotechnology Research Center; Yayoi 1-1-1, Bunkyo-ku, Tokyo
  113-8657, Japan (E-mail:uhaba@mail.ecc.u-tokyo.ac.jp,
  Tel:81-3-5841-3070, Fax:81-3-5841-8030)
  Location/Qualifiers
    1..1110
    /organism="Spingomonas sp. KAI"
    /mol_type="genomic DNA"
    /strain="KAI"
    /db_xref="taxon:164608"
    1..876
    /gene="carAa"
    <1..876
    /gene="carBa"
    /codon_start=1
    /transl_table=11
    /product="putative oxygenase component of carbazole
    1,9a-dioxygenase"
    /protein_id="BAB88910.1"
    /db_xref="GI:20135961"
    /translation="GDPYHGYRWDDRLVDILITNPGSVQIGRRALKTPPVEAKG
    LIFVYVGGEPPLIEDVYPGLDENRAIHGHRILVASNWLGAENGFDAGHVFHNK
    SILVKGNDIILPLGAPGDPDLTSEVAAGKPKGVYDLGHSVPVPEGMIEGKPAI
    HNGTSGKVAISISLWLVKVPWPPDLTQFEMVYVVDVETSHLFPQILGKVTSK
    RGLASFREFHKEWGLAALNGFNDDILARESMEFFYADDRGWSSEILFEPDRAIIW
    EKLASQHRNGQEAR"
    876..1109
    /gene="carBa"
    876..>1109
    /gene="carBa"
    /codon_start=1
    /transl_table=11
    /product="putative small subunit of 2'-aminobiphenyl
    2,3-diol"
    /protein_id="BAB88911.1"
    /db_xref="GI:20135962"

```

```

/translation="MGTTASYPVNRMLQBELFTNPGNVLFRADREALYERYGLSSAQR
AALDEGGFGLTAVGLHPVLQMHFMLTNPAPA"
ORIGIN
Query Match 75.8%; Score 859.2; DB 1; Length 1110;
Best local Similarity 99.7%; Pred. No. 3.5e-186;
Matches 861; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 274 TATCAGGCTGGACATATCGCTGGGACGATGGCGCTCGTGATATCCTCACAACCC 333
Db 13 TATCAGGCTGGACCTATCGCTGGGACGATGGCGCTCGTGATATCCTCACAACCC 72
QY 334 GGCAGTGTGAGATCGCGCGCGCTTTGAAGACGCTTCCCGGTTGAAGAGGCAAGGT 393
Db 73 GGCAGTGTGAGATCGCGCGCGCTTTGAAGACGCTTCCCGGTTGAAGAGGCAAGGT 132
QY 394 CTTATCTTCTGTTTACGTAGCGGCGGACCAACCGCTTATCAAGATGTGCCGCC 453
Db 133 CTTATCTTCTGTTTACGTAGCGGCGGACCAACCGCTTATCAAGATGTGCCGCC 192
QY 454 GGCCTTCTTGATGAAAACCGCGCATTCACGGCCAAACATCGGCTCGTGGCTCGAACTGG 513
Db 193 GGCCTTCTTGATGAAAACCGCGCATTCACGGCCAAACATCGGCTCGTGGCTCGAACTGG 252
QY 514 CGCTTGGGTGCGAAAACCGCTTTGATGGGGCAGCTTCTTCAATTCACAGATTCGATC 573
Db 253 CGCTTGGGTGCGAAAACCGCTTTGATGGGGCAGCTTCTTCAATTCACAGATTCGATC 312
QY 574 CTGCTGAAGGGCAACGATATCATCTTCCGCTTGGCTTTGGCGCTCGCGATCCCGACAG 633
Db 313 CTGCTGAAGGGCAACGATATCATCTTCCGCTTGGCTTTGGCGCTCGCGATCCCGACAG 372
QY 634 CTTACGGCTTCCGAGTTCGCGGCAAGCCCAAGAGGTGTTTACGATCTGCTTGGCGAG 693
Db 373 CTTACGGCTTCCGAGTTCGCGGCAAGCCCAAGAGGTGTTTACGATCTGCTTGGCGAG 432
QY 694 CATTGCTGCGGCTTTTCGAAGCATGATCGAAGGAAACCTGCAATCCATGGCAACATT 753
Db 433 CATTGCTGCGGCTTTTCGAAGCATGATCGAAGGAAACCTGCAATCCATGGCAACATT 492
QY 754 GGCAGCAAGCGCTCGCCATCAGCATATCGATCTGCTGCGGGCGCTACTCAAGGTGGA 813
Db 493 GGCAGCAAGCGCTCGCCATCAGCATATCGATCTGCTGCGGGCGCTACTCAAGGTGGA 552
QY 814 CCGTGGCGGATCCCGAGCTCAGCAGTTCGATGCTGCGCTCGATGAGACAGCAG 873
Db 553 CCGTGGCGGATCCCGAGCTCAGCAGTTCGATGCTGCGCTCGATGAGACAGCAG 612
QY 874 CACCTTCTTCCAGACGCTGGCAAGTCTGACCTCAAGAAACGCGGAGACTCCTTC 933
Db 613 CACCTTCTTCCAGACGCTGGCAAGTCTGACCTCAAGAAACGCGGAGACTCCTTC 672
QY 934 GAGCGAGATTCACCGAANAATGGGTAGGCTCGCGCTTAAAGGCTTCAATGATGACGAC 993
Db 673 GAGCGAGATTCACCGAANAATGGGTAGGCTCGCGCTTAAAGGCTTCAATGATGACGAC 732
QY 994 ATCATGGCAGTGAATCGATGAGCGCTTCTACGCTGATGATCGCGGTGGTCCGAAGAA 1053
Db 733 ATCTTGGCAGTGAATCGATGAGCGCTTCTACGCTGATGATCGCGGTGGTCCGAAGAA 792
QY 1054 ATCTGTTCCAGCGGACCGCAATCATCGAGTGGCGGGGCTTCCAGTCAAGCAAT 1113
Db 793 ATCTGTTCCAGCGGACCGCAATCATCGAGTGGCGGGGCTTCCAGTCAAGCAAT 852
QY 1114 CGCGGATTCAGGAAGCAGCTTGA 1137
Db 853 CGCGGATTCAGGAAGCAGCTTGA 876

RESULT 4
AB095952
LOCUS
DEFINITION
  Janthinobacterium sp. J3 car operons, partial and complete cds.
  20017 bp DNA linear BCT 04-FEB-2003
  AB095952

```

```

ACCESSION      AB095952
VERSION        AB095952.1  GI:28201193
KEYWORDS
SOURCE
ORGANISM      .
               Janthinobacterium sp. J3
               Janthinobacterium sp. J3
               Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
               Oxalobacteraceae; Janthinobacterium.
REFERENCE
AUTHORS      1. Wadada, J., Nojiri, H., Nakai, S., Inoue, K., Shintani, M., Yoshida, T.,
               Habe, H. and Omori, T.
               Molecular diversity of carbazole-utilizing bacteria and their
               catabolic genes
               Unpublished
JOURNAL
REFERENCE
AUTHORS      2. Inoue, K., Nojiri, H., Nakai, S., Endoh, T., Urata, M., Ashikawa, Y.,
               Saiki, Y., Yoshida, T., Habe, H. and Omori, T.
               Divergent structures of carbazole-degrading car operons isolated
               from several gram-negative bacteria
               Unpublished
JOURNAL
REFERENCE
AUTHORS      3. (bases 1 to 20017)
               Nojiri, H. and Inoue, K.
               Direct Submission
JOURNAL
TITLE        Submitted (11-NOV-2002) Hideaki Nojiri, Biotechnology Research
               Center, The University of Tokyo; 1-1-1, Yayoi, Bunkyo-ku, Tokyo
               113-8657, Japan (E-mail:anojiri@mail.ecc.u-tokyo.ac.jp,
               Tel:81-3-5841-3067, Fax:81-3-5841-8030)
FEATURES
source
1. 20017
   /organism="Janthinobacterium sp. J3"
   /mol_type="genomic DNA"
   /strain="J3"
   /db_xref="taxon:213804"
   /country="Japan"
   /count="609"
   /gene="ORFU1"
   /gene="ORFU1"
   /count="609"
   /gene="ORFU1"
   /function="unknown"
   /codon_start=1
   /transl_table=11
   /product="hypothetical protein"
   /protein_id="BAC56729.1"
   /db_xref="GI:28201194"
   /translation="MEHRIFCTVPKVKANKPRAYNRDNCUNYARQDFPQEGAAKAL
   QIGTFRRNRKGRMTFPOSQPNKDTSYPRQYETGAKTLTIQSHCKEFGMLTY
   SAEKSKVQOASAKCSIVQSAALCSIVQHCAA"
   985..1752
   /gene="ORFU2"
   /count="1752"
   /gene="ORFU2"
   /function="putative transfer protein"
   /codon_start=1
   /transl_table=11
   /product="transfer protein"
   /protein_id="BAC56730.1"
   /db_xref="GI:28201195"
   /translation="MRPWFRCSEVIKEREMHTVHFVLMKGVGKMSASLLAQLI
   ERALYCATDPTNMTFAHYKALNVEHFNISDNAKYTRKFDALIIKRAFESDCV
   VDTGASFLEMLYHLENKIIDLLESTGRKVLVHTPLVGGQAMDETIRGLQSILEFFS
   APVVWNVYFGVKDGVGFKETSLYQNAKGTGLGVFLQELTAETSGODVANWAKK
   YLTFDEVLSAPETDLVTRHRLGIVKXHIPDLAKIDL"
   complement(2511..3413)
   /gene="ORFU3"
   /count="3413"
   /gene="ORFU3"
   /function="putative transposase ORFB protein"
   /codon_start=1
   /transl_table=11
   /product="transposase"
   /protein_id="BAC56731.1"
   /db_xref="GI:28201196"
   /translation="MRAFIQYCHAYGVEPICKMVQVAPSGYWRVYAAQORNALRCAR
   VQRDDVLSVIERVQANLVYQADKVRQLRRREGTEVARCTVERLRRKAGLGVNRG
   KVTVTAAADTKAPCLDRVNRQNAQPNQLWVSDFTYVSTWQGFVTVAFVIVFAARR
   IVGWRVSCNRTDFVLDAEQALYARQPERNELVHSDRGSGYVSIKYSRLABAGIE
   PSVSGKGSYDNLAAETINGLYKAEILHRRAPWKTRAEVLALEWVSWFNRHRLLEP
   LGYIPPAEAEANYVKQLSSQAIPA"
   complement(3410..3733)
   /gene="ORFU4"
   /count="3733"
   /gene="ORFU4"
   /function="putative transposase ORFA protein"
   /codon_start=1
   /transl_table=11
   /product="transposase"
   /protein_id="BAC56732.1"
   /db_xref="GI:28201197"
   /translation="MKQPKISPEVIERAVRLVSEASEYSSQWAAIESIAAKIGCTP
   ETLRWYRQQRDGTGQRPPTAEERIKTLEREVRELRKANEILLASAFFAQLG
   RHFPK"
   3924..4253
   /gene="ORFU5"
   /count="4253"
   /gene="ORFU5"
   /function="putative DNA-binding protein"
   /codon_start=1
   /transl_table=11
   /product="DNA-binding protein"
   /protein_id="BAC56733.1"
   /db_xref="GI:28201198"
   /translation="MANIDLSGYNLAEKLGLOHDIEREIKSRQHDLOKAREQILALA
   QEAGVSVEELLAGGKSKTGKQVARYQNPSDNSQVTWGRGRQPKTAELGAKGK
   ALDDPRI"
   4765..5106
   /gene="ORFU6"
   /count="5106"
   /gene="ORFU6"
   /function="unknown"
   /codon_start=1
   /transl_table=11
   /product="hypothetical protein"
   /protein_id="BAC56734.1"
   /db_xref="GI:28201199"
   /translation="MVGPAAWYKVMNMRNHTIAALLVATLACNRSRPGLEGTK
   ADGLVPMITIFRPGTEAMGVIEKVDTTNGNVVYKTKDGLMKGSAIAFTVNRNTA
   NNPHYTLHRIR"
   5647..6171
   /gene="ORFU7"
   /count="6171"
   /gene="ORFU7"
   /function="putative cold shock protein"
   /codon_start=1
   /transl_table=11
   /product="cold shock protein"
   /protein_id="BAC56735.1"
   /db_xref="GI:28201200"
   /translation="MRLQGRVIEWNDERGFGFIQNGTSTERTFLHISALRKTGRPLL
   GSLVYETRRDDKGRLOAAELTGKLSQSSGSGFSALVGLVLLIMFGYGVYVRF
   SNOGTSASVYKIVLARDALRPHPEFQCTPEKSKCSKASCAELPHQERCNVPKMD
   GDGIGICEECWCN"
   6306..6818
   /gene="ORFU8"
   /count="6818"
   /gene="ORFU8"
   /function="unknown"
   /codon_start=1
   /transl_table=11
   /product="hypothetical protein"
   /protein_id="BAC56736.1"
   /db_xref="GI:28201201"
   /translation="MKIFASSLALILATAVPSTAQVATPASEARPTLLRALDGSWRM
   SGVDRGQVITYNSGAALQCTFTMEMKDVQPAQYEVVIFGDKASQTVIAHMD
   NGAKYISPHGTGQITGNSIQLLFPYQSGNFRNTWTVYHGTSSWTFLVESAQPDGSK
   HPARYEVCRN"
   6950..7378
   /gene="ORFU9"
   /count="7378"

```

CDS
6950..7378
/gene="ORF09"
/function="putative transposase ORFA protein"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC56737.1"
/db_xref="GI:28201202"
/translation="MARALLPDLWSLIAHLPAAHRRSPKGRPRIDRAAPTGLFV
LKTGPMEYLPRLGCGSGMTCCWRLLHEMLLAGVYQRIHETILRLREHDIWDRAS
VDAASVPAPAGEHTGRNPTDRGLKCKHLLWMSVGFHW"
7348..7629
/gene="ORF10"
7348..7629
/gene="ORF10"
/function="putative transposase"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC56738.1"
/db_xref="GI:28201203"
/translation="MVDRGPPLVAQISGAQVHDSRELIPLIVESVPVAVKGLAGRKR
FGKLHADRAYASRAHRAWRRRGIAARIYGVSRERLGRWVWVET"
complement(7862..8617)
/gene="CarR"
complement(7862..8617)
/gene="CarR"
/function="putative transcriptional regulator of car
operon"
- /codon_start=1
/transl_table=11
/product="regulatory protein"
/protein_id="BAC56739.1"
/db_xref="GI:28201204"
/translation="WDNTSNYLDQAPKASARAVKETSEHIVERQSRQTVRIAKA
IANGLLGRFAPGRVLEADLTKEIGSRNPLREAFGRVLAEGLLQVEPPRGAIVAKL
SRKELVDLSVRECLSEWAARLAENIDKSPFRNYLEASLENPSTILPENYSDYLE
DNRFHSLILKLAGNHLPLKLDLQFPFQSFAPFRMYDANIYQNSTAEHYEILQAIL
AGDAADASTRAAHVHRTSKLLYSLPDHFFNSD"
complement(8620..8880)
/gene="ORF12"
complement(8620..8880)
/gene="ORF12"
/function="unknown"
- /codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC56740.1"
/db_xref="GI:28201205"
/translation="WICQVEFLDNIALNIPPCDGLLSFEBQASVCYPRKIALQNC
SNTSFYFRYQNFLLIKFLKDMYRMTVWSNAHLVKNKLGK"
9141..10130
gene

Query Match 37.2%; Score 422.6; DB 1; Length 20017;
Best Local Similarity 62.8%; Pred. No. 3.5e-86;
Matches 691; Conservative 0; Mismatches 404; Indels 6; Gaps 2;
QY 31 AGAACCAAGTTTGGAGCCTTATATCCGTCGAAACTTCGGTTCCGAAACCATTTGGTAT 90
DB 10244 AGAGTAAGGGCTGGCGCCCTACGTGTGATCGAAGCTAGGCTTTCCGAATCATTTGGTAC 10303
QY 91 CCCGTTCCCTCGCGAGGAATCGCGAAGTACTCCCGTCCCGTCAAGCTCTCGGA 150
DB 10304 CCGGTGATGTTTTCGAGGAGATCAACGAGGGCGAGCGAGAGCTAAACTCTCGGT 10363
QY 151 GAGAAATTCCTCAATCGCGTGGGGCGCAAGTCTATGCGATCCAGGACAGTGGCTG 210
DB 10364 GAGAACTTCCTCGTCAATCGTATCGATGGGAGCTGATTTGCCCTCAAAGACCGCTGCTG 10423
QY 211 CATCGCGGTGTAACGCTTTCGACCGGGTCCAGTGTCTATTCAGAACACCATATCTGCG 270
DB 10424 CATCGCGGTGTCAGTCTGTCTGTCAAAAGTCAGTGCAGAAAGCAAGTGCAGATCATGC 10483

QY 271 TGGTATCACGGCTGGACATATCGTGGGACGATGGCCGCTCGTGCATATCTCTCACAAC 330
DB 10484 TGGTATCACGGCTGGACCTATCGTGGGAAAGCGGCTTCTGTGTGACATCTTGACGAAT 10543
QY 331 CCCGGCAGTGTGCAGATCGCGCGCGCTTTGAGACGTTCCCGGTTGAAAGGCGCAA 390
DB 10544 CCACGAGCGCACAGATCGGTGCACAGAAGCTGAAAACCTATCCAGTGCAGGAAGCCAG 10603
QY 391 GGTCTTATCTTCTGTTTACGTAGCGGACCGGCAACCAACCGCCCTTATCGAAGATGTGCGG 450
DB 10604 GGTGCGTCTTCTTATCTTCTGCGATGCGGACCCGCTCCCTTGGCCCGCGATACACCA 10663
QY 451 CCCGGCTTCTTGCATGAAAACCGCGCATTCAGGGCAACATCGGCTCGTGGCTCGCAAC 510
DB 10664 CCAATTTCTTGACATGACATGGAATCTTAGGGAAGACCAATATATCAAGTCTAAC 10723
QY 511 TGGCGCTTGGGTGCGGAAAACGCTTTGATGGGGGACGCTTTCATTCACAAGAAATTCG 570
DB 10724 TGGCGCTCGCTGTGGAACGCTTTGATCGAGCCACATTTATATTCACAAGACTCG 10783
QY 571 ATCTGCTGAAGGGCAACGATATCATCTTGGCGCTTGGCTTTGGCGC---TGGCGATGCC 627
DB 10784 ATTCTGCTCAAGGACACGATCTTGGCTTGCACACTAGGTTTTCGCGCAGCGCGCATCGA 10843
QY 628 GACCACTTACGGT--TCCGAGGTTGCTGGGGCAAGCCCAAGAGTGTGTACCATCTG 684
DB 10844 AGCAACAAACGCTGTGCTGACGATGATCTGCTGGCGCAAGGCTGTGTACGACCTA 10903
QY 685 CTTGGCGAGATTTCGGTGGCGGTTTTCGAAAGCATGATCGAAGGCAAACTGCAATCCAT 744
DB 10904 ATTGGCGAATCGGGTTCAGTGTTCGAGGGAACCTATCGGGGGCAAGTGTCCGCGAA 10963
QY 745 GGCACATTTGGGCAAGCGCTCGCCATCAGCATATCGATCTGGCTGCGGGCTACTC 804
DB 10964 GTGCTTACGGCGAATAAATTGTAGCGAAGCATATCTCCATTTGGCTCCAGGTGTTCTC 11023
QY 805 AAGTTCGAACCGTGGCGGATCCGAGCTCAAGCATGTCGAATGTTAGTGGCTGCGGTCGAT 864
DB 11024 AAGTCAATCCGTTCCCAATCCGACATGATGTCAGTTTCAATGTTAGTGGCTGCGGATGAT 11083
QY 865 GAGACCGACCTCTCTACTTCCAGACGCTGGGCAAGTGTGACGTCAAGAGAGCGGCA 924
DB 11084 GAAATACTCACTATTACTTCCAAACTCTTGGCAACCAATGTGCCAATGACGAAGACGG 11143
QY 925 GACTCTTCGAGCGAGAAATTCACGAAAATGGGTAGGCTCGCGCTTAACGGCTTCAAT 984
DB 11144 AAGAAGTACGACGACAGAGTTTCGAAGTAAAGTGAAGCCGATGCGCTCGAAGGATTCAC 11203
QY 985 GATGACGACATATGCGACGCTGAATGATGAGAGCGGTTCTACCTGATGATCGCGTTGG 1044
DB 11204 AACGACGACATCTGGGCGCGGAAAGCAATGGTGGGATTTCTACGCGCATGATAAGGCTGG 11263
QY 1045 TCCGAAGAAATCCTGTTTCGAGCGGACCGCAATCATCGAGTGGCGGGGCTTGCACGT 1104
DB 11264 GTCACGAGATTTGTTTCGAGTCCGACGAGGCTATCGTGGCATGCGCAAGTTGGCGAGC 11323
QY 1105 CAGCACAATCGCGGATTCAG 1125
DB 11324 GAACACACAGGATTTTCAG 11344

RESULT 5
LOCUS E17292
DEFINITION gDNA encoding oxidase which converts dibenzo-p-dioxin to a diphenyl ether triol.
ACCESSION E17292
VERSION E17292.1 GI:5711975
KEYWORDS JP 1998257895-A/1.
SOURCE Pseudomonas sp.
ORGANISM Bacteria; Proteobacteria.
REFERENCE 1 (bases 1 to 1161)

AUTHORS Omori, T. and Takami, K.
 TITLE OXIDASE GENE ORIGINATING FROM MICROORGANISM AND REMOVAL OF DIOXIN
 JOURNAL WITH THE SAME
 COMMENT Patent: JP 1998257895-A 1 29-SEP-1998;
 ASAHI CHEM IND CO LTD
 OS Pseudomonas sp.
 PN JP 1998257895-A/1
 PD 29-SEP-1998
 PF 18-MAR-1997 JP 1997084401
 PI OMORI TOSHIO, TAKAMI KAZUTAKA
 PC C12N15/09, A62D3/00, B09C1/10, C02F3/34, C07H21/04, PC
 C12N1/21//C12N9/02,
 CC (C12N15/09, C12R1/38), (C12N1/21, C12R1/19), (C12N9/02, C12R1/19);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC Key
 FH Location/Qualifiers
 FT source 1..1161
 FT /organism='Pseudomonas sp.'
 FT /strain='CA10'.
 FT Location/Qualifiers
 FT 1..1161
 FT /organism='Pseudomonas sp.'
 FT /mol_type='genomic DNA'
 FT /db_xref='taxon:306'

ORIGIN

Query Match 36.9%; Score 419.4; DB 6; Length 1161;
 Best Local Similarity 62.6%; Pred. No. 2e-85;
 Matches 689; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Qy 31 AGAACCAAGTTGGGACGCTTATATCCGTGCGAAACTCGGTTCCGAAACCACTTGGTAT 90
 Db 37 AGAGTAAAGCTGGGCGCCTACGTGATGCGAAGCTAGGCTTTCGCAATCATTGGTAC 96

Qy 91 CCGGTCGCTCGGAGCGAATCGCGAAGGTAAGTCTCCGTTCCGTCAGTCTCTGGGA 150
 Db 97 CCGGTGATGTTTCGAAGAGATCGAGGCGGCGAGCGAAGACATTAAGCTCTCGGT 156

Qy 151 GAGAAGATCTGTCATCGCGTGGGCGCAAGTCTATCGATCCAGGACAGTGCCTG 210
 Db 157 GAGAATCTGCTCGTCAATCGTATCGATGGAAGCTGTATTGCCCTCAAGGACCGTGCCTG 216

Qy 211 CATCGGCTGTAAGCTTTCGACCGGCTCGAGTCTATTCGAAGAACCATATCTCTGC 270
 Db 217 CATCGGCGCTCAGTTGTTCGGTCAAAGTCAGTCAAGCAAGAAAGTCGACGATCATGC 276

Qy 271 TGTATCACGCGCTGGACATATCGTGGGACGATGGCGGCTCGTTCGATATCCTCAAAAC 330
 Db 277 TGTATCACGCGTGGACCTATCGTGGGAGACGGGCTTCTGTGCGACATCTTGAGTAT 336

Qy 331 CCGGCGAGTGTGAGATCGCGCGGCGGCTTTGAAGACGTTTCCCGGTGAAGAGGCCAAA 390
 Db 337 CCGACAGCGCAGATCGGTGACAAAGCTGAAAGCTTAAAGCTTACCCAGTGCAGGAAGCAAG 396

Qy 391 GGTCTTATCTTCTGTTACGTAGCGAGCGGCGAAGCAAGCGGCTTATCGAAGATGCGCG 450
 Db 397 GGCTGCTCTTCATTATCTTGGGATGAGGACCTCTCTCTCTTGGCCCGGATAGCCCA 456

Qy 451 CCGGCTTCTTGTATGAAGAACCGGCGCATTCAGGCGCAACATCGGCTCGTGGCTCGAAC 510
 Db 457 CCCAAATTCCTTGACGATGATGGAATCTCTCGGGAAGAACCAATCATCAAGTCTAAC 516

Qy 511 TGGCGCTTGGTGGGAAAGCGCTTGTATGCGGGGCGAGTCTTCAATTCACAGATTGCG 570
 Db 517 TGGCGCTGCTGTGGAAACGCTTTCGATCCGAGCCACATTTATTCACAAAGACTCA 576

Qy 571 ATCTGGTGAAGGCAACGATATCATCTGCCCTTGGCTTTGCGCC---TTCGCGATCC 627
 Db 577 ATTCTGTCGAAGGCAACGATCTTGGCTTGGCCACTAGTTTTCGCGCAGAGGGATCGA 636

Qy 628 GACGAGCTTACGCGT---TCCGAGGTTCTCGCGGCAAGCCCAAGGTGTTTACGATCTG 684
 Db 637 AAGCAACAACTCGTGTGTTGACGATGACGTCGTGCGACGCAAGGGTGTTTACGATCTT 696

Qy 685 CTTGGCGAGCATTCGGTGCCTGTTTTCGAAGGCAATGATCGAAGCAACACCTGCAATCCAT 744
 Db 697 ATTGGCGAACATGGGTCCCAAGTCTTTGAGGGAATCTATCGGGGGGAGTGTGTCGCGAA 756

Qy 745 GCGAACATTTGCGAGCAAGCGCTGCGCATCAGCATATCGATCTGGCTGCGGGGCTACTC 804
 Db 757 GGTGCTTACGGCGCAAAATTTGTAGCGAACGATATCTCCATTTGGCTCCCGGGTGTCTC 816

Qy 805 AAGTTCGAACCGTGGCGGATCCCGAGCTCAGCAGTTCGAATGCTAGTCCCGTCCGAT 864
 Db 817 AAGTTCATCGTTCCTCCCAATCCGACATGATGATGCTGAGTGTGATGCTCCCATTTGAC 876

Qy 865 GAGACCAACCTCTTACTTCCAGACGCTGGCGCAAGTCTGTCAGCTCAAGGAAGCGCA 924
 Db 877 GAAACACACACTATTACTTCCAAACTCTTGGCAACCATGTGCCAATGACGGAACCG 936

Qy 925 GACTCTTCGAGCGAGATTCACGAAATGGGTAGGCTCGGCTTAAGGCTTCAAT 984
 Db 937 AAGAATTACGAACAGAGTTTCGAAGCAAGTGGAAACCGATGGCGCTCGAAGGATCAAC 996

Qy 985 GATGACGACATCATGGCACGTTGAATCGATGAGCGCTTCTACGCTGATGATCGCGTTGG 1044
 Db 997 AACGATGACATCTGGGCTCGCAAGCTATGTTGATTTCTACGCCGATGATAAAGGCTGG 1056

Qy 1045 TCCGAGGAATCTCTTCCAGCGCGGACCGGCAATCATCGATGCGGGGCTTCCAGT 1104
 Db 1057 GTCACAGAGATTTGTTTCGAGTGGAGGCTATCTGTCGATGCGCAAGCTCGCGAGC 1116

Qy 1105 CAGCACATCGCGCATTCAG 1125
 Db 1117 GAACACATCAGGTATTTCAG 1137

RESULT 6
 E17293
 LOCUS
 DEFINITION gDNA encoding oxidase which converts dibenzo-p-dioxin to a diphenyl ether triol.
 ACCESSION E17293
 VERSION E17293.1 GI:5711976
 KEYWORDS JP 1998257895-A/2.
 SOURCE Pseudomonas sp.
 ORGANISM Bacteria; Proteobacteria.
 REFERENCE 1 (bases 1 to 1200)
 AUTHORS Omori, T. and Takami, K.
 TITLE OXIDASE GENE ORIGINATING FROM MICROORGANISM AND REMOVAL OF DIOXIN WITH THE SAME
 JOURNAL Patent: JP 1998257895-A 2 29-SEP-1998;
 COMMENT ASAHI CHEM IND CO LTD
 OS Pseudomonas sp.
 PN JP 1998257895-A/2
 PD 29-SEP-1998
 PF 18-MAR-1997 JP 1997084401
 PI OMORI TOSHIO, TAKAMI KAZUTAKA
 PC C12N15/09, A62D3/00, B09C1/10, C02F3/34, C07H21/04, PC
 C12N1/21//C12N9/02,
 CC (C12N15/09, C12R1/38), (C12N1/21, C12R1/19), (C12N9/02, C12R1/19);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC Key
 FH Location/Qualifiers
 FT source 1..1200
 FT /organism='Pseudomonas sp.'
 FT /strain='CA10'.
 FT Location/Qualifiers
 FT 1..1200
 source

/organism="Pseudomonas sp."
/mol_type="genomic DNA"
/db_xref="taxon:306"

ORIGIN

Query Match 36.9%; Score 419.4; DB 6; Length 1200;
Best Local Similarity 62.6%; Pred. No. 2e-85;
Matches 689; Conservative 0; Mismatches 406; Indels 6; Gaps 2;
QY 31 AGAACCAAGGTTTCGGAGCCTTATATCGGTGCGAACTCGGTTCCGAAACCACTTGGTAT 90
DB 94 AGACTAAAGCTGGCGCCCTACGTGATGCGAGCTAGGCTTTCGCAATCATTTGGTAC 153
QY 91 CCCGTTCCCTCGCGAGCGAAATCGCCAGAGTACTCCCGTTCCCGTCAAGCTCCTGGGA 150
DB 154 CCGGTGATGTTTTCGAAGAGATCGACGAGCGCGAGACACTAAACTGCTCGGT 213
QY 151 GAGAAGATTCTGCTCAATCGGTGGCGGCAAGTCTATGCGATCCAGGACAGTGGCTG 210
DB 214 GAGAACTTGTCTGCTCAATCGTATCGATGGGAGCTGTATGCTCAAGACCGCTGCTG 273
QY 211 CATCGGCTGTAACGCTTTCCGACCGGCTGAGTGTATTTCCAAAGACACATATCTGC 270
DB 274 CATCGCGCGTCCAGTTGTCTGCTCAAAAGTCTGAGTCAAAACGAAATCGACATCACA 333
QY 271 TGGTATCAGCGCTCGACATATCGTGGAGAGATGCGCGCTCGTCAATATCCTCAAAAC 330
DB 334 TGGTATCAGCGCTCGACATATCGTGGAGAGATGCGCGCTCGTCAATATCCTCAAAAC 333
QY 331 CCCGCGAGTGTGAGATCGCGCGCGCTTTGAAGACGTTTCCGCTTGAAGAGCGCAAA 390
DB 394 CCGCAAGCGCACAGATCGGTGCGCAAAAGCTGAAACTTACCAGTGCAGGAGCCCAAG 453
QY 391 GGTCTTATCTCTGTTACGTAGGCGAGCGGCAACCAACGCGCTTATCGAAGTGTGCG 450
DB 454 GCGTGGCTTCTATTTATCTTGGCGATGGGACCTCTCTCCCTTGGCCCGCATACCCA 513
QY 451 CCCGCTTCTTGTGATAAAGCGCGCATTCACGCGCAACATCGGCTCGTGGCTCGAAC 510
DB 514 CCCAATTTCTTGTGAGATGACATGAAATCTCTCGGAGAACCAATCATCAAGTCTAAC 573
QY 511 TGGCGCTTGGTGGGAAAAAGCGTTTGTATCGCGGGGACAGTCTTCTATCAAGAAATCG 570
DB 574 TGGCGCTTGGTGGGAAAAAGCGTTTGTATCGCGGGGACAGTCTTCTATCAAGAAATCG 633
QY 571 ATCTGTTGAAGGCAACAGATATCATCTGCGCTTGGCTTTGGCGC---TGGCGATCCC 627
DB 634 ATCTGTTGAAGGCAACAGATATCATCTGCGCTTGGCTTTGGCGC---TGGCGATCCC 693
QY 628 GACGAGTTACGCT---TCCGAGTTGCTCGCGGCAAGCCCAAGGTGTTTACGATCTG 684
DB 694 AAGCAACAACTCGTGTGTTGACGATGACCTGCTCGGACGCAAGGGTGTATTACGATCTT 753
QY 685 CTTCGCGAGCATTCGGTGGCGGTTTTCGAAGGATGATCGAAGCAACCTGCAATCCAT 744
DB 754 ATTGGGCAACATGGGTCCCGAGTGTGAGGAACTATCGGGGCGAAGTGTGTCGCGAA 813
QY 745 GCGCAACATTCGACAAAGCGGTGCCATCAGCATATCGATCTGCTGCTCGCGGCGTACTC 804
DB 814 GGTGCTACGCGCAAAATTTGACGCAACATATCTCCATTTGCTCCCGGGTGTCTC 873
QY 805 AAGTTCGAACCGTGGCGGATCCCGAGCTCAGCAGTTCGATGATGTTGCTGCGCTGAT 864
DB 874 AAGTTCGAACCGTGGCGGATCCCGAGCTCAGCAGTTCGATGATGTTGCTGCGCTGAT 933
QY 865 GAGACCAACCACTCTACTTCCAGACGCTGGCGCAAGTCTGAGCTCAAGAGGAGCGGCA 924
DB 934 GAAACACACACTATTACTTCCAACTCTTGGCAACCATGTGCAATGACGAGGACGG 993
QY 925 GACTCTTCGAGCAGAAATTCAGAAATATGGGTAGCGCTCGCGCTTAACGGCTTCAAT 984
DB 994 AAGAATTACGAAACAGAGTTTCGAAAGCAAGTGGTGAACCGATGGCGCTCGAAGGATTCAC 1053

QY 985 GATGACGACATCATGCGACGATGAATCGATGGAGCGGTTCTACGCTGATGATCGCGTTGG 1044
DB 1054 AACGATGACATCTGGGCTCGGAGCTATGGTGGATTTCTACGCGGATGATAAGGCTGG 1113
QY 1045 TCGAAGAAATCTGTTTCGAGCCGACCGCGCAATCATCGAGTGGCGGGGCTTCCAGT 1104
DB 1114 GTCAACGAGATTTTTCGAGGTGACGAGGCTATCGTGGCATGCGCAAGCTGCGGAGC 1173
QY 1105 CAGCAATTCGCGCATTCAG 1125
DB 1174 GAACATCAGGCTATTACAG 1194

RESULT 7
AB001723
LOCUS
DEFINITION
Pseudomonas stutzeri carbazole catabolic genes, complete cds,
clone:PSB185.
AB001723
ACCESSION
AB001723.1 GI:3293057
VERSION
KEYWORDS
carbazole dioxygenase subunit; ferredoxin reductase; unknown;
ferredoxin; 2-hydroxy-6-oxo-6-(2'-aminophenyl)hexa-2,4-dienoate
hydrolase; 2'-aminobiphenyl-2,3-diol 1,2-dioxygenase.
SOURCE
Pseudomonas stutzeri
ORGANISM
Pseudomonas stutzeri
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (sites)
Ochiyama, N., Miyachi, S. and Omori, T.
Cloning and nucleotide sequence of carbazole catabolic genes from
Pseudomonas stutzeri strain OM1, isolated from activated sludge
J. Gen. Appl. Microbiol. 44, 57-63 (1998)
2 (bases 1 to 6880)
Ochiyama, N.
Direct Submission
Submitted (10-MAR-1997) Naoki Ochiyama, Chemicals Inspection &
Testing Institute, Kurume Research Laboratories; Chuo-machi 19-14,
Kurume, Fukuoka 830, Japan (E-mail: KYS04304@niftyserve.or.jp.
Tel: 0942-34-1500, Fax: 0942-39-6804)
Location/Qualifiers
1..6880
/organism="Pseudomonas stutzeri"
/mol_type="genomic DNA"
/strain="OM1"
/db_xref="taxon:316"
/clone="psb185"
571..1725
/codon_start=1
/transl_table=11
/product="carbazole dioxygenase subunit"
/protein_id="BAA31266.1"
/db_xref="GI:3293058"
/translation="MANVDEAILKRVKGMAPYVDALGFRNHYPVWFSKEIDGEPK
TLKLGLENLVNRIDGKLYCLKRCLHGVOLSVKVECKTKSTITCWYHMYRWEDG
VLCDILNPTSAIQGRQKLYPVOEAKGVFIYLGDDPPLARDTPNFDLDDMEI
LGNQILKSNRRLAVENGDFSHYIHKDSILVNDLALPLGFPAGGDRKQOTRVD
DVVGKRGYDLIGEHGVFVETGGTIGGVRREGAYGEKIVANDISILWPLVKVNPFP
NPMNQFWYPIVDENTHYFQTLGRCANDEERKNYQBFESKPKWALGFFNNDI
WAREAMVDFYADDKGWNVEILFEVDEAIVAWRKLASEHNQIGTQAHVSG"
1837..2991
/codon_start=1
/transl_table=11
/product="carbazole dioxygenase subunit"
/protein_id="BAA31267.1"
/db_xref="GI:3293059"
/translation="MANVDEAILKRVKGMAPYVDALGFRNHYPVWFSKEIDGEPK
TLKLGLENLVNRIDGKLYCLKRCLHGVOLSVKVECKTKSTITCWYHMYRWEDG
VLCDILNPTSAIQGRQKLYPVOEAKGVFIYLGDDPPLARDTPNFDLDDMEI
LGNQILKSNRRLAVENGDFSHYIHKDSILVNDLALPLGFPAGGDRKQOTRVD
DVVGKRGYDLIGEHGVFVETGGTIGGVRREGAYGEKIVANDISILWPLVKVNPFP
NPMNQFWYPIVDENTHYFQTLGRCANDEERKNYQBFESKPKWALGFFNNDI
WAREAMVDFYADDKGWNVEILFEVDEAIVAWRKLASEHNQIGTQAHVSG"
3049..3321

CDS
CDS
CDS

```

/codon_start=1
/transl_table=11
/product="2'-aminobiphenyl-2,3-diol 1,2-dioxygenase"
/protein_id="BAA31268.1"
/db_xref="GI:3293060"
/translation="MARVEVDRLIQDMSKKEGLGRVIDTPSDVFEYGLTPPERTAL
LEGTPQALASIGVHPILQMHVLYMKNPEMATHYSIKDYSMDLKGGA"
3321..4130
/codon_start=1
/transl_table=11
/product="2'-aminobiphenyl-2,3-diol 1,2-dioxygenase"
/protein_id="BAA31269.1"
/db_xref="GI:3293061"
/translation="MGKIVAGGTSHTILMSPKGCESAARVNVNGIAELGRRLKEARPD
VLVITSDHNMNINLSQPFREVGADSYTPMGDMIDPRDLVPGSEVGRALAQADE
DGFOLCQAEVSLDHGIMPIFLVGMKEIPVVPVIVNINTDPIPSARRCVALAESIRQ
AIERKTPQCEVRVAVAGGSHWLCVPRHGEVSEKFDHVMDELARNGNAEELVAMGNE
AIIIDGGNAGVEIILTWINAAVASEASSEGEKVFYEAMTQWFTGIGGVEHFVK"
4120..4992
/codon_start=1
/transl_table=11
/product="2-hydroxy-6-oxo-6-(2'-aminophenyl)hexa-2,
4-dienoate hydrolase"
/protein_id="BAA31270.1"
/db_xref="GI:3293062"
/translation="MLNKAKEQISKSBSAYVERVFNAGVETRYLEAGQGPVILING
GGAGSESGNWRNVIPILARHYVIAVMDLGFGKTAPDI EYTDREIRHLHDRIKAM
NFDGKSVGTVSGMSGATGLGVSLVHSLVNALVLMGSAGLVIEIHEDLPIINVDYFTR
EGMHLVYALINDYGIKIDDAMINSRYIYADEATRKAYVATMONTREQGLFIDYDPFI
RKVPFTVTVHGKODKQVVPVETAYTKFLDLIDDSGWYIIPCHGHWAMIEHPEDFANATL
SFLSRADITRAAA"
5013..5336
/codon_start=1
/transl_table=11
/product="ferredoxin"
/protein_id="BAA31271.1"
/db_xref="GI:3293063"
/translation="MNQILVKVCAASDQFQTIIRVNRVGAAPLAVYRVGDQFYATED
TCTHGIASLSGTLGDGVIECPFHGGAFNVCTGMPASSPCTVPLGVFEVEKEGEVTV
AVEKK"
5370..5687
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="BAA31272.1"
/db_xref="GI:3293064"
/translation="WADLSVITERVTVKAVGENSLDAVVKFDPPEGVTHIDGWSIPN
RVSNEDLPSDITIKIKLENPEKILNQDLGPWALATGRFELRGDIRIATRLDKVFGIA
PSM"
5717..6706
/codon_start=1
/transl_table=11
/product="ferredoxin reductase"
/protein_id="BAA31273.1"
/db_xref="GI:3293065"
/translation="WYQLKTEGQAPGTCGSKSLINLSALANGIGFPYECASGGCGVCKR
FELLEGNVQSWPAPGLSGSDREKGNRHACQCVALSDIRIKVAQVDKVPITPISR
MEAVEVEVYRTHDLLSVLRDTPANFLPQGFCLVEAQLPGVWVAYSANLKNPESJ
IWEFYIKRVPTGCRSPMLFENRKEGABLFITPGMTGFFRPGTGRSLCTIGGAGLSY
AAALASMRSDTGVKFLFYGSRTPRADWIDIDIDDKLEVQVQWMEDTLWQGP
TGFHOVDVAALLBTLPEYETLYIAGPPMDVATVMLLHGKVDPRDOIHPDPAP"

```

ORIGIN

[illegible]

RESULT 8
D89064

LOCUS	D89064	6881 bp	DNA	linear	BCT 12-SEP-1997				
DEFINITION	Pseudomonas sp. gene for carbazole 1,9a-dioxygenase, meta-cleavage enzyme, meta-cleavage compound hydrolase, complete cds.								
ACCESSION	D89064								
VERSION	D89064.1	GI:2317677							
KEYWORDS	reductase component of carbazole 1,9a-dioxygenase; ferredoxin component of carbazole 1,9a-dioxygenase; meta-cleavage compound of hydrolase; meta-cleavage enzyme; terminal dioxygenase component of carbazole 1,9a-dioxygenase.								
SOURCE	Pseudomonas sp.								
ORGANISM	Pseudomonas sp.								
REFERENCE	Bacteria; Proteobacteria.								
AUTHORS	Sato,S.I., Ouchiama,N., Kimura,T., Nojiri,H., Yamane,H. and Omori,T.								
TITLE	Cloning of genes involved in carbazole degradation of Pseudomonas sp. strain CA10: nucleotide sequences of genes and characterization of meta-cleavage enzymes and hydrolase								
JOURNAL	J. Bacteriol.	179 (15),	4841-4849	(1997)					
MEDLINE	97386424								
PUBMED	9244273								
AUTHORS	Sato,S.I., Nam,J.W., Kasuga,K., Nojiri,H., Yamane,H. and Omori,T.								
TITLE	1,9a-dioxygenase in Pseudomonas sp. strain CA10								
JOURNAL	J. Bacteriol.	179 (15),	4850-4858	(1997)					
MEDLINE	97386425								
PUBMED	9244274								
REFERENCE	3 (bases 1 to 6881)								
AUTHORS	Omori,T.								
TITLE	Direct Submission								
JOURNAL	Submitted (14-NOV-1996) Toshio Omori, The University of Tokyo, Biotechnology Research Center; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113, Japan (E-mail:aseigv@hongo.ecc.u-tokyo.ac.jp, Tel:03-3812-2111(ex.3067), Fax:03-5802-3326)								
FEATURES	Location/Qualifiers								
source	1..6881								
	/organism="Pseudomonas sp."								
	/mol_type="genomic DNA"								
	/strain="CA10"								
	/db_xref="taxon:306"								
gene	572..1726								
	/gene="carAa"								
CDS	572..1726								
	/gene="carAa"								
	/codon_start=1								
	/transl_table=11								
	/product="terminal dioxygenase component of carbazole								
	1,9a-dioxygenase"								
	/protein_id="BAA21728.1"								
	/db_xref="GI:2317678"								
	/translation="MANVDEAILKRVKGMAYVDKLGFRNHYTVMFSEKIDGEPK								
	TLKLLGENLLVNRIDGKLYCLKDRCLHRGVLSVKVECKTKSTITCYHWATYRWEDG								
	VLCDLTNPTAQIGROKLYTPVOEAKGCVFVLGDGPPPLARDTPPNFLDDMEI								
	LGKNOI IKSNWLA VENGFDPSHIYHKDSILVKNDDIALPLGAPGDRKQQRVVD								
	DDVGRKGYDLIGHGVFVEGTIGGEVREGAYGEKIANDISILWPGVLKVNPPF								
	NPDMMQFVYFIDENTHYFOTLGKPCANDERKNYEQEFESKWKPMALGFGFNDDI								
	WAREAMVDYADDKGNWNEILFEVDEAIVAMRKLASEHNQIGIQTAHVSG"								
	1838..2392								
gene	/gene="carAb"								
CDS	1838..2392								
	/gene="carAb"								
	/codon_start=1								
	/transl_table=11								
	/product="terminal dioxygenase component of carbazole								
	1,9a-dioxygenase"								
	/protein_id="BAA21729.1"								
	/db_xref="GI:2317679"								
	/translation="MANVDEAILKRVKGMAYVDKLGFRNHYTVMFSEKIDGEPK								
	TLKLLGENLLVNRIDGKLYCLKDRCLHRGVLSVKVECKTKSTITCYHWATYRWEDG								
	VLCDLTNPTAQIGROKLYTPVOEAKGCVFVLGDGPPPLARDTPPNFLDDMEI								
	LGKNOI IKSNWLA VENGFDPSHIYHKDSILVKNDDIALPLGAPGDRKQQRVVD								
	DDVGRKGYDLIGHGVFVEGTIGGEVREGAYGEKIANDISILWPGVLKVNPPF								
	NPDMMQFVYFIDENTHYFOTLGKPCANDERKNYEQEFESKWKPMALGFGFNDDI								
	WAREAMVDYADDKGNWNEILFEVDEAIVAMRKLASEHNQIGIQTAHVSG"								
	1838..2392								
gene	/gene="carAb"								
CDS	1838..2392								
	/gene="carAb"								
	/codon_start=1								
	/transl_table=11								
	/product="terminal dioxygenase component of carbazole								
	1,9a-dioxygenase"								
	/protein_id="BAA21730.1"								
	/db_xref="GI:2317680"								
	/translation="MARYEVDRLIQDMSKKEGLIGRVIDTPSDPPEYGLTTPPERTAL								
	LESTPOLASIGVHPILQMRYLKYKNPEVATHVSIKDISMLKGG"								
	3322..4131								
gene	/gene="carBb"								
CDS	3322..4131								
	/gene="carBb"								
	/codon_start=1								
	/transl_table=11								
	/product="catalytic subunit of meta-cleavage enzyme"								
	/protein_id="BAA21731.1"								
	/db_xref="GI:2317681"								
	/translation="MGKIVAAGTSHILSPKCEESAAARVNVNGTAEALGRRLKEARPD								
	VLVIITSDHFNINLSMPQRFVVGDIADSPMDMDIPRDLVPGSGEVGATLAQADE								
	DGFDLQAEYSLDHGMIPILFMGMKEIPVVPVI VNIIDPIPSARRCVLALESIRQ								
	AIEKTPDSCRVAVVGAGLSHLVCPVPHGEVSEKEDHVMDELVRGNAEKLAVMGNE								
	AIIQQGNAGVEILLTWIMAAVASEASGEKVFYEANTQMTGILGGVEFHVK"								
	5014..5337								
gene	/gene="carAc"								
CDS	5014..5337								
	/gene="carAc"								
	/codon_start=1								
	/transl_table=11								
	/product="ferredoxin component of carbazole								
	1,9a-dioxygenase"								
	/protein_id="BAA21733.1"								
	/db_xref="GI:2317683"								
	/translation="NNQIWLKVCASADMOPGIRRVNRVGAAPLAVYRVGDOPYATED								
	TCTHGIASLSEGLDGVIECPFHGAFNVCTGMPASSPTVPLGVFEVKEGEVYV								
	AGEKK"								
	5718..6707								
gene	/gene="carAd"								
CDS	5718..6707								
	/gene="carAd"								
	/codon_start=1								
	/transl_table=11								
	/product="ferredoxin reductase component of carbazole								
	1,9a-dioxygenase"								
	/protein_id="BAA21735.1"								
	/db_xref="GI:2317685"								
	/translation="MYQLKIEGQAPQTCGSKSLVLSALANGIFPYECASGCGVCK								
	FELLEGNVSMWPFADPLGSLRSRDRKGNHLAGCVALSDLRKVAQKVPYPTISR								
	MEAEVVEVALTHLLSVLRITDGPANFLPGQCLVEAQLPGVAVAYSMANLKNPEG								
	INFEYTKRYPTGTFSPMLPENKEGARFLFTGPMGTFPFGTGRKSLCIGGGAGLSY								
	AAAIASASRETDKPKVLPGYSGSTRDAVRWIDIDIDEDKLVVQAVTSDLSLWGP								
	TGFHVQVDAALLETUPEYIYLAGPPWDAVTRVMLLKGVRDQIHLHDAFF"								
ORIGIN									
Query Match	36.9%; Score 419.4; DB 1; Length 6881;								
Best Local Similarity	62.6%; Pred. No. 1.9e-85;								
Matches	689;	Conservative	0;	Mismatches	406;				
				Indels	6;				
				Gaps	2;				
QY	31	AGAACCAAGTTTGGGAGCCTTATATCCGTGCGAAACTCGGTTCCGAAACCACTGGTAT	90						
Db	602	AGAGTAAAGGCTGGGCGCCCTACGTGGATCGAAGTAGGCTTTCGCAATCAATGGTAC	661						
QY	91	CCCGTTCGCTCCGAGCGAATAACGCGGAGGTAACCTCCGCTCCGCTCAAGCTCCCTGGGA	150						
Db	662	CCCGTGTATGTTTCGAAAGANGATCGAGGCGCGACCGAAGACATAAACTGCTCGGT	721						
QY	151	GAGAGATTCCTCAATCGCTGGCGCGCAAGGTCTATCGCATCCAGACAGGTCCCTG	210						
Db	722	GAGAACTTGTCTGCTCAATCGTATCGATGGAAGCTGATTGCTCAAGGACCGCTGCTG	781						

```
Qy 211 CATCGCGGTGTAACGCTTTTCGACCGGGTCGAGTGTCTATTTCCAAAGAACACCATATCTCTGC 270
|||
Db 782 CATCGCGCGTCCAGTTGTCTGGTCAAAAGTCGAGTGCAGAAACGAAAGTCGAGTACATGTC 841
|||
Qy 271 TGGTATCACGCGTGGACATATCGTGTGGAGCATGTGGCGCTCTGTCGATATCTCCACAAAC 330
|||
Db 842 TGGTATCACGCGTGGACATATCGTGTGGAGCATGTGGCGCTCTGTCGATATCTCCACAAAC 901
|||
Qy 331 CCAGGAGTGTGAGATCGCGCGCGCTTTGAAGAGTGTCCCGTGTGAAGAGGCAAA 390
|||
Db 902 CCAGGAGGCGACAGATCGGTGCAAAAGTGTGAAGAGTGTCCCGTGTGAAGAGGCAAA 961
|||
Qy 391 GGTCTTATCTGTTTACGTAGGAGCGGCGACCAACCGCGCTTATCGAGATGTCCG 450
|||
Db 962 GGTCTTATCTGTTTACGTAGGAGCGGCGACCAACCGCGCTTATCGAGATGTCCG 1021
|||
Qy 451 CCAGGCTTCTGTGATGAAGAGCGCGCTTACCGGCAACATCGGTGTGGGCTCGAAC 510
|||
Db 1022 CCAGGCTTCTGTGATGAAGAGCGCGCTTACCGGCAACATCGGTGTGGGCTCGAAC 1081
|||
Qy 511 TGGCGCTTGGTGGGAACCGCTTTGATGCGGGGACGCTTCTATTCACAGATTCG 570
|||
Db 1082 TGGCGCTTGGTGGGAACCGCTTTGATGCGGGGACGCTTCTATTCACAGATTCG 1141
|||
Qy 571 ATCTGTGTGAAGGCAACGATATCTGCGCTTGGCTTGGCC---TGGCGATCCC 627
|||
Db 1142 ATCTGTGTGAAGGCAACGATATCTGCGCTTGGCTTGGCC---TGGCGATCCC 1201
|||
Qy 628 GACGAGCTTAGCGT---TCGAGGTTGTGCGGGGCAAGCGCTTGTACGATCTG 684
|||
Db 1202 AAGCAACAACTCGTGTGTGAGCATGACGCTGTCGGAGCGCAAGGCTGTGTACGATCTT 1261
|||
Qy 685 CTTCGCGAGCATTCGCTGCGCGTTTTCGAGGCGATGATCGAAGCAACCTGCAATCCAT 744
|||
Db 1262 ATTCGCGAGCATTCGCTGCGCGTTTTCGAGGCGATGATCGAAGCAACCTGCAATCCAT 1321
|||
Qy 745 GCGCAACATGCGCAGCAACGCGCTGCGCATACGATATCGATCTGGCTGCGCGGCGTACTC 804
|||
Db 1322 GGTGCTTACGCGCAAAATTTGTAGCGAAGCATATCTCCATTTGGCTCCCGGCTGTTCTC 1381
|||
Qy 805 AAGTGTGAACGCTGCGGATCCCGAGTTCAGGCTTTCGAGTTCGATGTCGCGGCTCGAT 864
|||
Db 1382 AAGTGTGAACGCTTCCCAATCCCGACATGATGAGTTCGATGTCGCGGCTCGATGATGAC 1441
|||
Qy 865 GAGACCGACCACTTACTTCCAGACGCTGGGCAAGTTCGATGATGATGATGATGATGATG 924
|||
Db 1442 GAAACACACACTTACTTCCAACTCTTGGCAACCATGTGCAATGACGAGGAACGG 1501
|||
Qy 925 GACTCTTCGAGCGAGATTCACGAAATGGGTAGCGCTTCGCGCTTAACGCTTCAAT 984
|||
Db 1502 AAGAAATACGAAACAGAGTTCGAAAGCAAGTGGAAACCGATGCGGCTCGAAGGATCAAC 1561
|||
Qy 985 GATGACGACATCATGCGACGCTGAATTCGATGAGCGGCTTCTACGCTGATGATGCGGTTGG 1044
|||
Db 1562 AACGATGACATCTGGCTTCGCGAGCTATGCTGGATTTCTACGCGGATGATAAAGGCTGG 1621
|||
Qy 1045 TCCGAGAAATCTGTTTCGAGCGGACCGCGCATCATCGAGTGGCGGGGCTTGCAGT 1104
|||
Db 1622 GTCAACGAGATTTGTTTCGAGGTGAGCGAGGCTATCGTGGCATGCGCAAGCTGGCGAGC 1681
|||
Qy 1105 CAGCACATTCGCGGCAATTCAG 1125
|||
Db 1682 GAACACATCAGGATTTTCAG 1702
|||
```

```
RESULT 9
AB047548
LOCUS AB047548 4266 bp DNA linear BCT 03-MAY-2002
DEFINITION Pseudomonas resinovorans strain CA10 DNA, car gene cluster flanking
region.
ACCESSION AB047548
VERSION AB047548.1 GI:13094152
```

KEYWORDS

SOURCE
ORGANISM
Pseudomonas resinovorans
Pseudomonas resinovorans
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE

AUTHORS
Sato, S.I., Ouchi, Yama, N., Kimura, T., Nojiri, H., Yamane, H. and
Omori, T.
TITLE
Cloning of genes involved in carbazole degradation of Pseudomonas
sp. strain CA10: nucleotide sequences of genes encoding carbazole
of meta-cleavage enzymes and hydrolase
J. Bacteriol. 179 (15), 4841-4849 (1997)

JOURNAL

MEDLINE
PUBMED
97386424
9244273

REFERENCE

AUTHORS
Sato, S.I., Nam, J.W., Kasuga, K., Nojiri, H., Yamane, H. and Omori, T.
TITLE
Identification and characterization of genes encoding carbazole
1,9a-dioxygenase in Pseudomonas sp. strain CA10
J. Bacteriol. 179 (15), 4850-4858 (1997)

JOURNAL

MEDLINE
PUBMED
97386425
9244274

REFERENCE

AUTHORS
Nojiri, H., Sekiguchi, H., Maeda, K., Urata, M., Nakai, S., Yoshida, T.,
Habe, H. and Omori, T.
TITLE
Genetic characterization and evolutionary implications of a car
gene cluster in the carbazole degrader Pseudomonas sp. strain CA10
J. Bacteriol. 183 (12), 3663-3679 (2001)

JOURNAL

MEDLINE
PUBMED
21264379
11371531

REFERENCE

AUTHORS
Nojiri, H., Omori, T. and Habe, H.
TITLE
Direct Submission
Submitted (21-AUG-2000) Hideaki Nojiri, The University of Tokyo,
Biotechnology Research Center, 1-1-1 Yayoi, Bunkyo-ku, Tokyo
113-8657, Japan (E-mail: anojiri@mail.ecc.u-tokyo.ac.jp,
Tel: 81-3-5841-3054, Fax: 81-3-5841-8030)

JOURNAL

MEDLINE
PUBMED
21264379
11371531

FEATURES

source

1. .44286
/organism="Pseudomonas resinovorans"
/mol_type="genomic DNA"
/strain="CA10"
/db_xref="taxon:53412"
complement(214..1929)
/gene="ORF31"
complement(214..1929)
/function="putative membrane spanning protein"
/codon_start=1
/transl_table=11
/product="putative ABC transporter subunit"
/protein_id="BAB32742.1"
/db_xref="GI:13094153"
/translation="MGVINLAHGFYVWVAFRCPTPRERLVLRVGGSGSARPYTR
CLSKPACAAATTAITSTRYMTLAVFTNELITLVFGRFPFAMPTPEWYSSFFVEVL
PGLMTPVSRLLFIAAGAAVAIAMLLINHLRGLMIRAGADHHEMVGALGVNIARLY
LVFVFGSVLGLAGPMAAPLLSVEIGMEKVLITTFVIVVGVGVRGALAGALLIG
MYDGLRAVLPOLLQDLPAEVSGLSGELISASAVIMAVVLVVRPGLLPARALRR
TCSRVRVLDLCLATFALLPLASLHPEPLVSLFTLALYGMASLIDLILGYGAM
ISFGRAAFGLGVVGVGTGHTSOALFWGDSNALLVWPLSLLYCALFAVUVGY
LCLRTSGVQFIMTLAFQMLFFVLSLSLGGDDGILLISERITLFGIDLPVQFY
LCGLVLAFLFCRLVRSFYGLKQSRERISIGLHPLRVLGAFVAGVGGG
LAGILWANYAMFVSPDMGAWHKSGLMAMVILGVGTLLGPVLAHVILGLEQCSACC
PSTGC"

gene

CDS

gene

CDS

complement(2085..3245)
/gene="ORF30"
complement(2085..3245)
/function="putative ABC transporter substrate binding
protein"
/codon_start=1
/transl_table=11
/product="putative ABC transporter subunit"
/protein_id="BAB32743.1"

```

/db_xref="GI:13094154"
/translation="MKQLVGTLLAPVWAPALRPAQDPVVRIGFTITLSTPAGYLGED
SRDQLAIDDEGGKGVPSLLVEDDGLQARAKQIIDRMSLDGI.SLYGTGVIFSNV
LAAVNTATQDGYISNTPSPSLAGCKPKFYFVASYQNDTIHEWAGVAANDLGYKK
VILAPNTQGRDAGEGRKTFKGEVTIYTKLQDLDFSELARVSLPAIDAFQFHP
GGAGINFAKQYSSGLNKSIPMVVPVFSMDERMLAAGDVAKGNIVTLWNPSTNPA
NQAQVAFNDKYKVPYVYAAQTVDTABLIQALKTVDGNLKDPEALASALRNVTFDS
TRGDFAFKGNHADVWLLRVREADANGKLVQVPVQTIKAKQVDSFAECAL"
gene
CDS
complement(3405..4385)
/gene="tnpA3"
complement(3405..4385)
/gene="tnpA3"
/feature="tnpA3"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB32744.1"
/db_xref="GI:13094155"
/translation="MKQMTFADAAYAGKQKTRKELFLIEMDRVVPWKGLIALIEPHY
PKGEGRPAYPLMAMLRVLLQNNFGYSDPAMEEALYETTLRQFAGLNLERIPDET
ILNFRLEKHELALGILAVINGYLGRGLSRQGTIVDATLINAPSTKNKQGRDP
EMHQYKGNQYIFGMAHIGADDESLVHSVVGTAANVADVQVDKLHGDENVVCA
AGYVGEKRPHEGREVWQVAAARRSTYKLDKRSALYKAKKIEKAKQVRAKVEHS
FRVYKRGYTKVFRGLAKNTAQLVTLFALSNLWMARRHILLNAGEVRL"
gene
CDS
complement(4543..5571)
/gene="antC"
complement(4543..5571)
/gene="antC"
/feature="antC"
/codon_start=1
/transl_table=11
/product="reductase component of anthranilate
1,2-dioxygenase"
/protein_id="BAB32745.1"
/db_xref="GI:13094156"
/translation="MSHKVAFSFDGKTLFPFVQSHETLLDAALRNGINPLDCREGV
GTCGRCEGSGYQDYDDALSAEDLRQKMLTCQTRVSSDAFYDFPASSLCNAP
CGELKLVQVQSPNTAILHLDAGADGQDLFPQGYARLQIPGTHGORSYSFAN
RPSNGNQFTIRLLPDGLMGNVIRERSLIGDEILLEAFYLRHVKPLFVAVAG
KPLSALQMDIEAERGCGHPVHLVYGVRAADLCETQRIAAVAERIPGPRLPFGSS
ATPARTWTKRGYITEHFDLAEFRDQALDMLVLCGPPPWVESIKQWLDTQALDQRLY
BKFTQSN"
gene
CDS
complement(5590..6081)
/gene="antB"
complement(5590..6081)
/gene="antB"
/feature="antB"
/codon_start=1
/transl_table=11
/product="small subunit of terminal oxygenase component of
anthranilate 1,2-dioxygenase"
/protein_id="BAB32746.1"
/db_xref="GI:13094157"
/translation="MNSOLQYRVEQFLYRNAEACDAQDWDVYLSMFSDDSEFPLPOWD
SEHEVTRQPKRAMSLIYVNRGGLEDVRIETGKSASTVPMPTLHLSNVRIAPLA
DELELVQNVHTLYRLQTAQPFGRATYRLRPGESWKIARKVLLINDTINSVLDF
YHL"
gene
CDS
complement(6084..7487)
/gene="antA"
complement(6084..7487)
/gene="antA"
/feature="antA"
/codon_start=1
/transl_table=11
/product="large subunit of terminal oxygenase component of
anthranilate 1,2-dioxygenase"
/protein_id="BAB32747.1"
/db_xref="GI:13094158"
/translation="MSSARSVEQWKNFIEGCLDFPPADGVFRIARDIFTEPDLFQLE
ELIFPKNIYACHSEIAKHPMTIRAGRQPMIITRDGEGQLNALINACORHGTITL
RVKGNQSTFTICFPFANCIKSDGLRVKAPAGEPEFGKATRGKAKRIEISYGTFL
ISLDVGDQDSQEFIDAKVFPDMMVAQSPTEGEVLVPGKSAITYDGNWKLQNGENGLD
GYHVSIVHYNVAVTQHRQVQNSKGLTSADTLDYSLKLGAGDKETDDGWFAPFHNGSV

```

```

LFSDMPNPVVRPOYATIMPRLLQEQYQGGKQAKQMMHFLRNINLYPSMFFLDQISSQURI
IREIANKTEINSICLVGVGSDRENIROFDFNFYSGMGTDDLVDFRQKGF
QARLEWDSISRGQKQWVGPSTANSETIGTSPVLITGTETHEGLYVQNGWQRFLE
GLARKAQOPLKLRV"
complement(7829..8809)
/gene="tnpA2"
complement(7829..8809)
/gene="tnpA2"
/feature="tnpA2"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB32748.1"
/db_xref="GI:13094159"
/translation="MKQMTFADAAYAGKQKTRKELFLIEMDRVVPWKGLIALIEPHY
PKGEGRPAYPLMAMLRVLLQNNFGYSDPAMEEALYETTLRQFAGLNLERIPDET
ILNFRLEKHELALGILAVINGYLGRGLSRQGTIVDATLINAPSTKNKQGRDP
EMHQYKGNQYIFGMAHIGADDESLVHSVVGTAANVADVQVDKLHGDENVVCA
AGYVGEKRPHEGREVWQVAAARRSTYKLDKRSALYKAKKIEKAKQVRAKVEHP
FRVYKRGYTKVFRGLAKNTAQLVTLFALSNLWMARRHILLNAGEVRL"
gene
CDS
complement(9062..10318)
/gene="ORF24"
complement(9062..10318)
/gene="ORF24"
/feature="ORF24"
/function="channel-forming protein"
/codon_start=1
/transl_table=11
/product="PhaK-like protein"
/protein_id="BAB32749.1"
/db_xref="GI:13094160"
/translation="MFOASWFRGMQTSALALMTPMAQADFVEDSQLALGRNFYI
DRDFKGNPNKSRVSGTQGFDLRPNSTYEGTQFALDASQYAVRLDGGGGRGPD
IIPYSTRQFOATDYGRGYTAKRVYKTELKYEHRPTLPVATYDSDQLVTVYHGT
LIESKVDKLTLTGSGFTSISGESSENSEKMLKNGPDYRSDGDLNFGSVAFTPA
LGITYYQGLEDIYRCHYLGATLADLGSYALKTDLRYFNKEDDGLYICIDNRSY
LMTURKAGSFLGYQRMGLDSTPTTLNGYAPQYLVNVAFAVFNKESKSWQRY
DYDFAGLGLPLGLMTRYMRGSDIDRGALSDNSESRNTYLSVYVQEGPLKLGVELR
NINVRHGNDFDENRLITTYTWKFW"
gene
CDS
complement(10639..11655)
/gene="ORF23"
complement(10639..11655)
/gene="ORF23"
/feature="ORF23"
/codon_start=1
/transl_table=11
/product="putative regulatory protein"
/protein_id="BAB32750.1"
/db_xref="GI:13094161"
/translation="NMSTSLDVHVRDIRIERYDLNAGAROMAGICGPHLLKTOYPNRI
QFHHSANVLKSMSTILGVIEYDVTGIEDVEHLNYSLSPLVGEQHLKDKGCQVT

```

```

Query Match 36.9%; Score 419.4; DB 1; Length 44266;
Best Local Similarity 62.6%; Pred. No. 1.9e-85;
Matches 689; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Qy 31 AGAACCAAGCTTTGGGAGCCTTATATCCGTCGCAAACTCGGTCGCAAAACCATTCGTAT 90
Db 28539 AGGTAAAGCTGGGCGCCCTACGTGGATGGGAAGCTAGGCTTCGCAATCATTCGTAT 28598

Qy 91 CCGCTTCGCTCGCGAGCGAAATCGCGAAGGTACTCCCGCTTCGCTCAAGCTTCCTGGGA 150
Db 28599 CCGGTGATGTTTTCGAAAGAGATCGACGAGGCGGACCGAAGACACTAAACTGCTCGGT 28658

Qy 151 GAGAGAGTTCTGCTCAATCGCTGGGCGCAAGGCTTATCGCATCCAGGACAGTGCCTG 210
Db 28659 GAGAACTTCTGCTCAATCGTATCGATGGGAAGCTTATTCGCTCAAGGACCGCTCGCTG 28718

Qy 211 CATCGCGGTCTAACGCTTTCGACCGGTCGAGTGTCTATTCAGAAACACCATATCCTGC 270
Db 28719 CATCGCGGCTCCAGTTGTCGTCAGATCGATGCAAAACGAGTCGACGATCATGC 28778

Qy 271 TGGTATCAGCGCTGGACATATCGCTGGGACGATGGCGCGCTCGCTGATATCTCACAAC 330
Db 28779 TGGTACCACGCTGGACCTATCGCTGGGAAGACGGGCTTCTGTCGACATCTTTGACGAA 28838

```

```
QY 331 CC CGCAGTGTGCAGATCGCGCGCGCGCTTTGAAGACGTTCCCGGTTGAAGAGCCAAA 390
D 28839 CCGACAAGCGCACAGATCGGTGCGACAAAGCTGAAACTTACCAGTGCAGGAAGCCAA 28898
QY 391 GGTCTTATCTCGTTTACGTAGGCGGCGGCGGAAACCAAGCGCGCTTATCGAGATGTCGG 450
D 28999 GGCTGCGCTTCACTTATCTTGGCGATGCGGACCTCTCTCCCTGGCCCGGATACGCCA 28958
QY 451 CCCGCTTCCCTGTATGATAAACCAGCGGCGCATTCACGCGCCAAACATCGGCTCGTGGCTCGAAC 510
D 28959 CCCAATTTCCCTTGACGTACATGAAATCTCTCGGGAAGAACCAATCATCAAGTCTAAC 29018
QY 511 TGGCGCTTGGGTGCGGAAACCGCTTGTATCGGCGGCGACGCTTCTCATTCACAGAATCG 570
D 29019 TGGCGCTTGGGTGCGGAAACCGCTTGTATCGGCGGCGACGCTTCTCATTCACAGAATCG 29078
QY 571 ATCTGTGTGAAGGCAACGATATCATCTGCGCGTCTGCTTGGCGC---TGGCGATCCC 627
D 29079 ATCTGTGTGAAGGCAACGATATCATCTGCGCGTCTGCTTGGCGC---TGGCGATCCC 29138
QY 628 GACGAGTACCGT---TCCGAGTGTCTGCGGCGAAGCCAAAGCTGTTTACGATCTG 684
D 29139 AAGCAACAACTCGTGTGTTGACGATGACGCTGCGAAGCAAGGCTGTTACGATCTT 29198
QY 685 CTTCGCGAGCATTCGTCGCGGTTTTCGAGGCGATGATCGAAGCAACCTGCAATCCAT 744
D 29199 ATTCGCGAGCATTCGTCGCGGTTTTCGAGGCGATGATCGAAGCAACCTGCAATCCAT 29258
QY 745 GCGAATATGCGCAGCAAGCGCTCGCCATCAGCATATCGATCTGCTGCGCGGCGTACTC 804
D 29259 GGTGCTACGCGGAAATAATGTTAGCGAATATCTCCATTTGCTCCCGGTTCTC 29318
QY 805 AAGTCGACCGTGGCGGATCCCGAGTTCAGGCTACGAGTTCGATGTTGTCGCGGCTGAT 864
D 29319 AAGTCGACCGTGGCGGATCCCGAGTTCAGGCTACGAGTTCGATGTTGTCGCGGCTGAT 29378
QY 865 GAGACCGACCTACTTCTCCAGACGCTGCGGCAAGTCTGAGCTCAAAAGGAAGCGCA 924
D 29379 GAAACACACACTATTACTTCCAACTCTTGGCAACCATGTGCAATGACGAGGAACG 29438
QY 925 GACTCTTCGAGCGAGATTCACGAAATGTTGAGCTCGCGCTTACGGCTTACGGTTCAT 984
D 29439 AAGAAATACGAAACAGAGTTCGAAGCAAGTGAACCGGATGCGGCTCGAAGGATTCAC 29498
QY 985 GATCAGCATCATGCGGACGCTGATTCGATGAGCGCTTCTACGCTGATGATCGCGGTGG 1044
D 29499 AACGATGACATCTGGGCTCGGAGCTATGTTGATTTCTACGCGGATGATAAGGCTGG 29558
QY 1045 TCCGAAGAAATCTGTTTCGAGCGCGGACCGCGCAATCATCGAGTGGCGGCGCTTGCAGT 1104
D 29559 GTCAACGAGATTTGTTTCGAGTGGGACGAGGCTATCGTGGCATGCGGCAAGCTGCGGAGC 29618
QY 1105 CAGCACAATCGCGCATTCAG 1125
D 29619 GAACACAATCAGGTTATTCAG 29639

RESULT 10
AB088420 199035 bp DNA circular BCT 03-MAY-2003
LOCUS AB088420
DEFINITION Pseudomonas resinovorans plasmid pCAR1, complete sequence.
ACCESSION AB088420
VERSION AB088420.1 GI:26106080
KEYWORDS
SOURCE Pseudomonas resinovorans
ORGANISM Pseudomonas resinovorans
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 Maeda,K., Nojiri,H., Shintani,M., Yoshida,T., Habe,H. and Omori,T.
Complete nucleotide sequence of carbazole/dioxin-degrading plasmid
pCAR1 in Pseudomonas resinovorans strain CA10 indicates its
```

```
JOURNAL J. Mol. Biol. 326 (1), 21-33 (2003)
MEDLINE 22435905
PUBMED 12547188
REFERENCE 2 (bases 1 to 199035)
AUTHORS Nojiri,H., Shintani,M. and Maeda,K.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-2002) Hideaki Nojiri, Biotechnology Research
Center, The University of Tokyo; 1-1-1 Yayoi, Bunkyo-ku, Tokyo
113-8657, Japan (E-mail:anojiri@mail.ecc.u-tokyo.ac.jp,
Tel:81-3-5841-3064, Fax:81-3-5841-8030)
FEATURES
Location/Qualifiers
1..199035
/organism="Pseudomonas resinovorans"
/mol_type="genomic DNA"
/db_xref="taxon:53412"
/plasmid="pCAR1"
join(185744..199035,1..59548)
repeat_region
transposon="Tn4676"
gene
complement(join(198367..199035,1..81))
CDS
/gene="ORF188"
/function="putative ABC-transporter subunit"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative ABC-transporter subunit"
/protein_id="BAC41521.1"
/db_xref="GI:26106270"
translation="MNPOLSVQALEKTFGAVKATNGASFDIAPGELHAIIGPNGAGKS
TLIAOAGEIRPKGRIFLNGODITDPAHORPLRLGLARFQVQLYPEFTLLENVAV
AARADGOTGELKPMASDQSLAPADYLRARVGLRACHAQSALSGEROLEIAL
ALAPASVLLDEPWAGNAGQESARMTQJLLELKGRIYALLVEHDMADVAFALADRTV
LVYGTLLTGSVEQRNDPVRRAAYLGEHA"
gene
complement(1108..1929)
CDS
/gene="ORF31"
/function="putative membrane spanning subunit"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative ABC transporter subunit"
/protein_id="BAC41521.1"
/db_xref="GI:26106081"
translation="MGVTLNLAHGFYMGAFCAVTSLSGSMFAGLAAYIGSALYV
LLIETSIWRFLYNRDHLQVLAFLAVFTNELITVLRGSPFPAMPTFWSSFEVL
PGLMTPVSRLLFIARAGAVATAMLLLNHRLMIRAGADHDHVMGALGNARLYT
LVFPGCVLCGLAGFMAAPLISVEIGMERKVLITTFVIVVGGVSGVAGALLIG
MYDGLGRAYIFQLDQLLPABVSGTSLGGILSASAYIVMAVLLVVRPEGLLPARA"
gene
complement(2084..3244)
CDS
/gene="ORF30"
/complement(2084..3244)
/gene="ORF30"
/evidence=not_experimental
/codon_start=1
/transl_table=11
/product="putative ABC transporter subunit"
/protein_id="BAC41522.1"
/db_xref="GI:26106082"
translation="MKKQVLTLLAASCLASLAACADPVRIGFTTTLTSGVLGED
SRDGFQLAIDGEGKLGVPVSLVVEDDGLQRAKQIIDMSLDGLISLYTGVFSNV
LAAVYATKDGFIISNTGPSLQKQKPHYFVASYQNDITHMAGVANDLUGYKK
WVILAPNTQGGRDALGFKRTFKVEITITKQLDPSVELARVSLAPDAIFQFHP
GGAGINFARQYSGSLNKSIIMVVFVSMERMLAATGDVAKMNIIVLWPGSTNPA
NGAFVAKNDKYKRVPTVYAQTVDLARLGAALKTVQDNLKDPALRSALRNVTFS
TGFDAFGKNOHAVIDVLLRVEADANGKLVQVPVQTITAKQVDSFAEAL"
repeat_unit
3255..3258
/rpt_type=direct
repeat_region
3259..4448
/insertion_seq="ISPre2"
repeat_unit
3259..3274
```

```
gene /rpt_type=inverted
complement (3404..4384)
CDS /gene="tnpA3"
complement (3404..4384)
/notes="transposase of ISPre2"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC41523.1"
/db_xref="GI:26106083"
/translation="MKOMTFADAEYAGKQTRKELFLIEMDRVVPWPKGLIALIEPHY.
PKGEGRPAYPLMAMLRVHLQNWFGYSDPAMEEALYETTLRQFAGLSLERIPDET
ILNFRLLERHELAAGILAVINGYDGLRSLRQGTIVDATLINAPSTKNKQGRDP
EMHQTCKGNQYFGMAHIQVDDSEGLVSHVSTAAVADVTQVQLLHGDENVGAD
AGTVGKPEHEGREVINQIAARSTYKLDKRSALYKAARKIEKAQVRAKVEHP
FRVIKQFGYTKVRFRLAKNTAQLVTLFALSNLMMARRHLLANAGEVRL"
repeat_unit 4433..4448
/rpt_type=inverted
repeat_unit 4444..4447
/rpt_type=direct
complement (4543..5568)
/gene="antC"
complement (4543..5568)
/gene="antC"
/notes="AntC"
/codon_start=1
/transl_table=11
/product="reductase component of anthranilate"
/protein_id="BAC41524.1"
/db_xref="GI:26106084"
/translation="MSHKVAFSPADGKTLFPVQSHEILLDAALNGINIPLDCEGV
CGTCQKSGSQSDYDDEALSDLRQKMLTCQTRVSSDAAPYDFASLSCNAP
GPRGLGIVRAVEQSPNTAILHLDAGDQQLDFPGQYARQLPGTHQDSYISFAN
PNSGNQLQFLRLDGLMSYIERSLIGDEILAEPLGAFYLRHVDPKPLVFAVAG
TGLSAFLGMLDEAGCGGHPVHLVYGVNADLCETORIAAVERIPGFRFTPVIS
DRPDWGHKRGYITTEHFDLAEFRDQALDMLVLCGPPMVESIKQLDQALDQTRYLE
KPTQSN"
gene complement (5587..6078)
CDS /gene="antB"
complement (5587..6078)
/gene="antB"
/notes="AntB"
/codon_start=1
/transl_table=11
/product="small subunit of terminal oxygenase component of
anthranilate 1,2-dioxygenase"
/protein_id="BAC41525.1"
/db_xref="GI:26106085"
/translation="MNSQLQYRVEQFLYRNABACDAQDWDYLSMFSDEPHLPQWD
SEHEYTRDPKRAMSLIYPNRRGLEDRVFRITGKSASTVPMPTLHLLSNVRIAPLA
DDELEVQVNWHTLYRLQTAEQFFGRTYLRPDGESWKIARKHLLNDTINSVLDF
YHL"
gene complement (6081..7484)
CDS /gene="antA"
complement (6081..7484)
/gene="antA"
/notes="AntA"
/codon_start=1
/transl_table=11
/product="large subunit of terminal oxygenase component of
anthranilate 1,2-dioxygenase"
/protein_id="BAC41526.1"
/db_xref="GI:26106086"
/translation="MSSARSVEQWKNFIBGCLDFRADGVFRIARDIFTEPDLFLEM
ELIFENKMLVACHESIANKHDFMTWRAGROPMLITDEGOLNALINACOREGTLT
RVGKNQSTFTCFPHAWCYKSDRLVKNVAPGEYEGFPKATRGKARIESYKGFVP
ISUDVQDQSLQFELGDAKVPFDMVAQSTGUEVLPGKSAITTDGNWKLQENGLD
SVSTVYXNYAVTQVROQVNSEKGLTSADLDYKLGAGKDETDGDFAFNGHSV
LFSDMPNPSVPGYATIMPRLQEQYCGQKAQWMMHRLNLIYPSMFFLDQISSQLRI
IRFIANNKTEINSFCVLGVGSEADRENRIQFDFPNVSGMTGPDLLVEFREAQKGF
QARLERSDISRGVKGWVGPTANSETIGISPLVLTGTEFTHREGLYVNGHWNQRFLE
GLARKAEQSQLKREV"
```

```
repeat_unit 7676..7679
/rpt_type=direct
repeat_region 7680..8871
insertion_seq=ISPre1"
7680..7695
/rpt_type=inverted
complement (7826..8806)
/gene="tnpA2"
complement (7826..8806)
/notes="transposase of ISPre1"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC41527.1"
/db_xref="GI:26106087"
/translation="MKOMTFADAEYAGKQTRKELFLIEMDRVVPWPKGLIALIEPHY
PKGEGRPAYPLMAMLRVHLQNWFGYSDPAMEEALYETTLRQFAGLSLERIPDET
ILNFRLLERHELAAGILAVINGYDGLRSLRQGTIVDATLINAPSTKNKQGRDP
EMHQTCKGNQYFGMAHIQVDDSEGLVSHVSTAAVADVTQVQLLHGDENVGAD
AGTVGKPEHEGREVINQIAARSTYKLDKRSALYKAARKIEKAQVRAKVEHP
FRVIKQFGYTKVRFRLAKNTAQLVTLFALSNLMMARRHLLANAGEVRL"
8856..8871
/rpt_type=inverted
repeat_unit 8872..8875
/rpt_type=direct
9059..10315
/gene="ORF24"
9059..10315
/gene="ORF24"
/notes="putative channel-forming protein"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="Phak-like protein"
/protein_id="BAC41528.1"
/db_xref="GI:26106088"
/translation="MFOASMTGSMOTASLAALMTPLAODFVEDSOLALGLNRFYI
DRDFKGNPNXSRVGSWTQGFDLRFNSGYTEGTLQFALDASQVAYRLDGGGEGPDS
IIPXTRRQOATDYGRRGTAKRYSIKELKYGEHRPTLPVATIDDRSQVITYHGT
LIESKVDKLTLTGRTFETISGRESSNEXKMYLFNFPDVRSDGLNFGASAYFTA
```

QY	31	AGAACCAGGTTTGGGAGCCTTATATCCGTCGGAACCTCGGTTCCGAAACCTTGGTAT	90
DB	28517	AGAGTAAAGGCTGGGCGCCCTACGTGGATCGAAGCTAGGCTTTCCGCAATCATTTGGTAC	28576
QY	91	CCCGTTCCGCTCGCGAGGGAATCGCGAAGGTACTCCCGTTCCCGTCAAGCTCCTCGGA	150
DB	28577	CGGTGATGTTTTCGAAGAGATCGACGAGGCGGCGAAGACACTAATACTGCTCGGT	28636
QY	151	GAGAAGATTGCTCAATCGCGTGGCGGCAAGGTCTATGCGATCCAGGACAGTGCCTG	210
DB	28637	GAGAACTTGCCTCGTCAATCGTATCGATGGAAGCTGATTGCTCCCAAGGACCGCTCGT	28696
QY	211	CATCGCGTGTAAACGCTTTCGACCGGTCGAGTCTATTCCAGAGAACACCATATCTGC	270
DB	28697	CATCGCGCGTCCAGTTGTCGTTCAAGTCCAGTGCAGAACGAGTGCAGATCAGATGC	28756
QY	271	TGGTATCAGCGCTGGACATATCGTGGGACGATCGCGCTCGTTCGATATCTCTCACAAC	330
DB	28757	TGGTACCACGCGTGGACCTATCGTGGGAGACGCGCTTCTGTGCGACATCTTGACCAAT	28816
QY	331	CCCGCAGTGTGCAGATCGCGCGCGCTTTGAGACGTTCCCGGTGAGAGGCCAAA	390
DB	28817	CCGCAAGCGCACAGATCGGTCGACAAAGCTGAAACTTACCCAGTGCAGAGGCCAAG	28876
QY	391	GGTCTTATCTTCTGTTTACGTAGGCGACGCGGCAACCAACGCGCTTATCGAAGATGTGCG	450
DB	28877	GGCTGCGTCTTTCATTTATCTTGGCGATGGCGACCTCTCTCTTGGCCGCGATACGCCA	28936

Query Match 36.9%; Score 419.4; DB 1; Length 199035;
Best Local Similarity 62.6%; Pred. No. 1.8e-85;
Matches 589; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Db	7444	GAGAACTTGTCTGCTCAATCGATCGATGGGAGCTGTATTGCTCCTCAGGACCGCTGCGCTG	7503
Qy	211	CATCGCGGTGTAAAGCTTTCCGACCGGGTCTGAGTGTCTATTCCAGAAACACATATCTGCTG	270
Db	7504	CATCGCGCGCTCCAGTTGTTCGGTCAAAGTCTGAGTGCAGAAACGAAAGTCTGATCAGATGC	7563
Qy	271	TGGTATCACCGCTGGACATATCGCTGGGACGATGCGCGCTCTGTCGATATCTCTCACAAC	330
Db	7564	TGGTATCACCGCTGGACATATCGCTGGGACGCGCTCTGTCGATATCTCTCACAAC	7623
Qy	331	CCCGGACGTGTGACAGATCGCGCGCGCTTTGAAGACGTTTCCCGTGTGAAGAGCGCCAAA	390
Db	7624	CCGACAAAGCGCACAGATCGGTTCGACAAAAGCTGAAAACTTACCACAGTGCAGAAAGCCAA	7683
Qy	391	GGTCTTATCTTCGTTTACGTAGGCGACGCGGACCAACCGCGCTTATCGAAGATGTCCCG	450
Db	7684	GGCTGCGTCTTCATTTATCTTTGGGATGGCGACCTCTCTCCCTTGGCCCGCGATACGCCA	7743
Qy	451	CCCGGCTCTCTTGTATGAAAAACCGGCGCCATTCACGCGCCAAACATCGGCTCTGGCTCGAAC	510
Db	7744	CCCAATTTCTTGACGATGACATGGAATCTCTCGGAGAACCAATCATCAAGTCTAAC	7803
Qy	511	TGGCGCTTGGGTGCGGAAAAGCGCTTTGATCGGGGACGCTTCATTCACAGAAATTCG	570
Db	7804	TGCGCGCTCGCTGTGGAAGACGGTTTCGATCCGAGCCACATTTATTTATTCACAAAGACTCA	7863
Qy	571	ATCTCTGGTGAAGGGCAACGATATCATTTCTGCGCTTGGCTTTTGCGCC---TGGCGATCCC	627
Db	7864	ATTCTGTGTAAGCAACGATCTTGCCCTTGCCACTAGTTTCGCGCCAGGAGGGATCGA	7923
Qy	628	GACCAAGTTTACCGT---TCCGAGTTGCTCGGGCAAGCCAAAGTGTTCATGATCTG	684
Db	7924	AAGCAACAAACTCGTGTGTTGACATGACGTCTCGGACGCAAGGGTGTTCAGATCTT	7983
Qy	685	CTTCGCGAGCATTCGCTGCGCGTTCGAAAGGATGATCGAAGCAACCTGCAATCCAT	744
Db	7984	ATTGCGCAACATGGGTCCCAGTGTTCGAGGAACATATCGGGGCGGAAGTGGTCCGGAA	8043
Qy	745	GGCAACATTTGCGACGAAACCGGCTCGCCATCAGCATATCGATCTGCTGCTCCCGG	797
Db	8044	GGTCCCTACGCGCAAAAATTGTAGCAACGATATCTCCATTGGCTCCCGG	8096
RESULT 12			
PPY12655			
LOCUS			
DEFINITION	P.putida oxoS, oxoO, oxoH genes.		
ACCESSION	Y12655		
VERSION	Y12655.1		
KEYWORDS	2-oxo-1,2-dihydroquinoline 8-monooxygenase; oxoH gene; oxoO gene; oxoS gene; oxygenase; putative hydrolase; putative regulatory protein.		
SOURCE			
ORGANISM	Pseudomonas putida		
REFERENCE			
AUTHORS	Rosche, B., Tshisuka, B., Hauer, B., Linsens, F. and Fetzner, S.		
TITLE	2-oxo-1,2-dihydroquinoline 8-monooxygenase: phylogenetic relationship to other multicomponent nonheme iron oxygenases		
JOURNAL	J. Bacteriol. 179 (11), 3549-3554 (1997)		
MEDLINE	97315225		
PUBMED	9171399		
REFERENCE			
AUTHORS	Fetzner, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-APR-1997) S. Fetzner, Universitaet Oldenburg, Fachbereich 7, Postfach 2503, D-26111, Oldenburg, FRG		
FEATURES			
source			
	1. 3203		
	/organism="Pseudomonas putida"		
	/mol_type="genomic DNA"		
	/strain="86"		

gene	PKTYLEDLK
CDS	4160..5140 /gene="ORF10K23L" 4160..5140 /gene="ORF10K23L" /function="transposase" /note="100% identity with TnpA1, TnpA2 and TnpA4 in strain Cal10" /codon_start=1 /transl_table=11 /product="transposase" /protein_id="BAC56724.1" /db_xref="GI:28201187" /translation="MKQMTFADAAYAGKKEKOTRKELEIEMDRVVPWKGLIALTEPHY PRGGRPAYPMAMLRVHLQWFGYSDPAMEEALYETTLRQFAGLSLERIPDET INLFRLEKHELAGILAVINGYLGDRGLSRQGTIVDAILINAPSSTKNGKGRDP EHWQTKGQNYFYGNKAHIGVDDESLVHSVGTAAADVTDVQDKLLHGDENWVGAD AGYGVGRKEPEHEGREVIWQIAARSTYKLLDKRSALYKAARKIEKAAQVRAKVEHP FVIRKQFYGVKVRFRGLAKNTAQLVTLFALSNLWARRHLNANAGEVRL" 5508..7211 /gene="ORF9K23L" 5508..7211 /gene="ORF9K23L" /function="unknown" /note="100% identity with ORF9 in strain Cal10" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAC56725.1" /db_xref="GI:28201188" /translation="MEKLYMRAGRQPMIITDRGGQNALINACQHRGTTLTRVGKG NOSTFTCPHAWCYSDGLRVKAPGEYEPGDKATRLGKARIESYGFVIFSLDV DGDSLOEFLDAKQVFFDMVAQSPTELEVLPGKSAYTDGNMKNLNGENGLDGVHS TVHNVYA"VQHQROQETEMKTSTPASMOTSGVKAHVHALLCTVLACETVNVAS PNPSPEVATIGRAETEMKTSTPASMOTSGVKAHVHALLCTVLACETVNVAS DLPTNLGTSFLDGPSPGGLYFTQHLQYRSGSHRLDQAGNRVPLRPTDLRLDVS TQVYLTNODILGGALFLDLPVWVDRKIDDLNANVLNDRGVGDITFGPAIQWAP VNGPRPYRAHVEFOFLPTGSDYDLKAINFGSNFNFYAGTVMNPNWTSWR LHLYMKNKTDPGSPGATSGVAGQVHANFTTBYALTPOFRGLNGYMLNQTDT KINGRDVSRREKWAIGPAGVFSLSPHDLFLNAYFEQDAENRPEGKRMQLRWVHHF K" 7294..8096 /gene="ORF4K23L" 7294..8096 /gene="ORF4K23L" /note="100% identity with partial CarAa in strain Cal10" /codon_start=1 /transl_table=11 /product="partial terminal oxygenase component of carbazole" /protein_id="BAC56726.1" /db_xref="GI:28201189" /translation="MANVDEAILKRVKGWAPYVDADKLGRNHWYPMFSEKIDEGEPK TLKLGLENLVNRIDGKLYCLDKRCLHRGVQSVKVECKTKSTITCYHWAMTYRWEDG VLCDIITNPSAQIGROKLTVPVQAKGVFLYLDGDPPLPARDTTPNFUDDMEI LGNQIJKWIDLVGEBDSHVIYHKDSILVXONDIALPLGFAPEGDEKQQTWVD DDVVGKGVYDLIGRGVFPFESTIGGEVVRGAYGEKIVANDISINLPG"
ORIGIN	
Query Match	24.8%; Score 281.8; DB 1; Length 8096;
Best Local Similarity	62.1%; Pred. No. 6.4e-54;
Matches	480; Conservative 0; Mismatches 287; Indels 6; Gaps 2;
Qy	31 AGAACCAAGTTGGAGCCTTATATCGTTCGCGAACTCGGGTTCCGAAACCATTTGGTAT 90
Db	7324 AGAGTAAAGGCTGGCGCCCTTACGTGTATGCGAAGCTAGGCTTTTCGAATTCATTTGTTAC 7383
Qy	91 CCGGTTCCCTCGCGAGCGAAATCGCGAAGCTACTCCCGTTCCGCTCAAGCTCTCCGGA 150
Db	7384 CCGGTGATGTTTTCGAAAGAGATCGACGAGCGCGAGCGCAAGACACTTAAACTGCTCGGT 7443
Qy	151 GAGAGAGATTCTGCTCAATCGCGTGGCGGCAAGGTCTATGCGATCCAGGACAGTGCTG 210


```

gene
complement(1000..1377)
/gene="SC00858"
/note="synonym: SCM2.11c"
CDS
complement(1000..1377)
/gene="SC00858"
/note="SCM2.11c, unknown, len: 125 aa."
/codon_start=1
/transl_table=11
/product="hypothetical protein SCM2.11c."
/protein_id="CAB65639.1"
/db_xref="GI:6689149"
/db_xref="SPTREMBL:O9RCV3"
/translation="MQEETARSALDMFISFNASDDTYTALLSQALTSDDVVFVWGLG
RSGIAEVRFDIRKHPAGTGMVRCSAVDPEWARIRWVFPPDGGPLAGTDV
VHLRSLIDQIVFAGDIEPAAS"
complement(1546..2055)
/gene="SC00859"
/note="synonym: SCM2.12c"
complement(1546..2055)
/gene="SC00859"
/note="SCM2.12c, possible membrane protein, len: 169 aa.
Contains possible hydrophobic membrane spanning region"
/codon_start=1
/transl_table=11
/product="putative membrane protein."
/protein_id="CAB65640.1"
/db_xref="GI:6689150"
/db_xref="SPTREMBL:O9RCV2"
/translation="MTTSWTFPRRLADTALAALAPICACVLGALGALAVWTATGNAG
TPAIGVTDARLFLPGRVETAAFFKINTNGAQRDLRVTSSEVPEGISSHRMT
AGGAHRPPTESLFPVPAEGTDLMSPLSSDVTVPAAARWQAGDLVPTLHFSGRNEV
LAVVVRPGS"
complement(2052..4340)
/gene="SC00860"
/note="synonym: SCM2.13c"
complement(2052..4340)
/gene="SC00860"
/note="SCM2.13c, probable cation-transporting ATPase, len:
762 aa. Highly similar to many other cation transporters
e.g. Enterococcus hirae SW:COPA.ENTHR(EMBL:L13292),
copper/potassium-transporting ATPase A (EC 3.6.1.36) (727-
identity in 749 aa overlap. Contains Prosite hits to
PS01047 Heavy-metal-associated domain and PS00154 E1-E2
ATPases phosphorylation site. Also contains Pfam matches
to entries PF00122 E1-E2 ATPase, E1-E2 ATPase and PF00403
HMA, Heavy-metal-associated domain as well as multiple
possible membrane spanning hydrophobic domains."
/codon_start=1
/transl_table=11
/product="probable cation-transporting ATPase"
/protein_id="CAB65641.1"
/db_xref="GI:6689151"
/db_xref="GOA:Q9RCV1"
/db_xref="SPTREMBL:O9RCV1"
/translation="MAGEPSTVSQEVTDLAVGMTCACTVRVERKLAKLDVGSVNL
ATGARVHHPEVLPQGLVAASQGYTALPOSVEERRGSDDDAGTAQERDLRT
TVALLAVFVLVSNVPAQWPNQWLCFLVLAAPVAVGAWPFHRAARLRHSTMTD
TLVSLGVAFSWSXALFLGGADPDLEPFSLPTASDGVAAHYLEAAVGVPLFVL
AGRYLEARARGTGALRALAEKVAEVARVDTGGRRIPIEIRLVGVQVVRPGERV
ATDGTVVGSSAVDLSLVTCSEPAEVA PGTA VIAGVNVGGLAVRATVAGADTFLA
RI THLVTEAAGAKRQARQAKVGVFVFLPLATLVLFGLGADAPQALITASVA
VLVAVPCALGATPTALMAATGRGALGVNLVSGPQALSGLRHIDVNLDKTGLTFSG
HNSVARTPNFGIGSEQVAVLAGAVEQSGEHPGQAVTAYARTNVTLSLPEVTDFA
ALPGVRGVREGLREVTAPDDELVPVLDENSGAESAAHTPVVVRVDVGTGELFV
RELREGRVAVVGDGVNDAAALAGADLGTAMGTGTDAAIGAADVTLVRGDDALADA
VRLSRSTLRTVNLWAFQNVVVTMPVAVVGLSPMPAAVMSVSSLLVGVNSLRLR
AWOPARTSRPSAPARGESLR"
complement(2133..3707)
/gene="SC00860"
/note="Pfam match to entry PF00122 E1-E2 ATPase, E1-E2
ATPase, score 529.80, E-value 1.9e-155"
misc_feature
complement(2997..3017)
/gene="SC00860"
/note="PS00154 E1-E2 ATPases phosphorylation site"
complement(4209..4295)
/gene="SC00860"
/note="Pfam match to entry PF00403 HMA,
Heavy-metal-associated domain, score 37.40, E-value
1.1e-08"
complement(4212..4295)
/gene="SC00860"
/note="PS01047 Heavy-metal-associated domain"
complement(4340..4549)
/gene="SC00861"
/note="synonym: SCM2.14c"
complement(4340..4549)
/gene="SC00861"
/note="SCM2.14c, putative secreted protein, len: 69 aa.
Rich in the amino acid Gly. Contains possible N-terminal
region signal peptide sequence"
/codon_start=1
/transl_table=11
/product="putative secreted protein"
/protein_id="CAB65642.1"
/db_xref="GI:6689152"
/db_xref="SPTREMBL:O9RCV0"
/translation="MQRIILPFPALCGFLVLLVLAFAASYAVGRGVGPVAPGMHGGGIT
QGGHGDGDTMEEDDMGGMHGGGH"
complement(4813..5472)
/gene="SC00862"
/note="synonym: SCM2.15c"
complement(4813..5472)
/gene="SC00862"
/note="SCM2.15c, possible integral membrane protein, len:
219 aa. Similar to a protein of unknown function from
Mycobacterium tuberculosis SW:Y970 MYCTU(EMBL:Z79700)
hypothetical 22.9 KD protein (210 aa), fasta scores opt:
224 z-score: 235.4 E(): 9.8e-06 30.5% identity in 213 aa
overlap. Contains possible membrane spanning hydrophobic
domains and a possible N-terminal signal sequence."
/codon_start=1
/transl_table=11
/product="putative integral membrane protein."
/protein_id="CAB65643.1"
/db_xref="GI:6689153"
/db_xref="GOA:Q9RCU9"
/db_xref="SPTREMBL:O9RCU9"
/translation="WIAAGLWILTLMEAPALYGLWLRVPTTGLTGRVGHLLHAA
NSVLMIAVAVWPGVDLAVAPQVVLPAAGALMFVAASLWFGERSRSGAVRAVPHALM
MGAWMVAVAMGTSGAGHGGTGCHGHEGHAASGSLASMSLTATGPSVASLLAV
ALTAIGLVMLARALDLARVAPLPAGGPAPAGTDTTAAALDPACHAVNALGNVMAFLF
A"
complement(5469..6926)
/gene="SC00863"
/note="synonym: SCM2.16c"
complement(5469..6926)
/gene="SC00863"
Query Match 4.3%; Score 48.8; DB 1; Length 298450;
Best Local Similarity 48.2%; Pred. No. 1.5;
Matches 137; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
Qy 721 ATCGAGCGCAACCTCGAATCCATCGGCACATTGGCAGCAGCGCGTCCGATCAGGATA 780
Db 97961 ATGGAGCGCAGTACGCCCTTCCCTGGTACTCAACACCGGCCCTCTTCTTCAACAAG 97902
Qy 781 TCGATCTGCCTGCCGGGGCTACTCAAGTCCGACCGTGGCCGGATCCCGAGCTCACGCG 840
Db 97901 TCCCTCTTCGAGAGGCCCGGACTCGACCCCGAGCAGCCGCGAGGAGCTACGAGGTC 97842
Qy 841 TTCGATGGTAGCTGCCGCTCGATGAGCAGCACCTCTTCTTCCAGACCTCGGCGAAA 900
Db 97841 TTCGAGCGCCCTCGAATCGCCGACCAAGACCGCGCGAGGTCCGCCACCTCGCCAAC 97782

```

QY 901 GTCTGACGCTCAAGAGAGCGGACAGATCTCTTCGAGCGAGAGAAATCCAGAAAAATGGGTA 960
Db 97781 GTGCCACCATCGAGGACTTCGGCGCTACGGCGTCCCGCTCATGAACAAGAGGGCACC 97722
QY 961 GGCCTCGGCTTAAGCGTTCAATGATGACACATCATGGCAGC 1004
Db 97721 GCCTTCGCTTCAACGACGCCAAGGTTGTCAGCTCTCGCCCG 97678

RESULT 15
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS AF429315 Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE 11694876
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source
1..125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
repeat_region complement(35581..35746)
/rpt_type=tandem
/rpt_unit="ctg"
gene complement(<36507..>36887)
/gene="JPH3"
/note="synonym: JP3"
mRNA complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
CDS complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translation="MSSSGRPNFDGSGYCGWEDGKAHGVCTGPKGQGEYTGWS
HGFEVLGYTPSGNTYGTQWAGRGHIGLESKGMVYKGEWTHGPKRGYVRECA
NGAKYEGTWSNGLQDGYGTETYSQ"

ORIGIN

Query Match 4.2%; Score 47.4; DB 9; Length 125020;
Best Local Similarity 11.5%; Pred. No. 3.1;
Matches 106; Conservative 374; Mismatches 439; Indels 5; Gaps 2;
QY 80 ACCATTGGTATCCCTTCGCTCGAGCGAAATCCGGAAGGTACTCCCTTCCTCCGTC 139
Db 16929 MSMTYKSSWGSWCCWGRRRRSXGKWKYSRGMSKSMRYTGGSKMRSMMCTSSCY 16988

QY 140 AGCTCTGGGAGAGAGATTTCTCTCAATCGGCTGGCGCGCAAGGTCTTATGGATCAGG 199
Db 16989 ASWCCNCCWSCCNRSCCCMCERSYCCMEYCCACKCYNSSYMTMSASYNRSRYSWRSK 17048
QY 200 ACAGGTGCTCGCATCGGCTGTAAAGCTTTCGACCGGTCGAGTCTATTCCAGAAACA 259
Db 17049 CWSRMCSSRSRSKCKSRGCGSGMGKGGKSGYGRTRKRSRKGKMKAKWMTYRRSRMR 17108
QY 260 CCATATCTCTGCTGTATACAGGCTGTGACATATCGGTGGAGCGATGCGCCCTCGTCGATA 319
Db 17109 KMYSSKGMVCMYCWGRRGCTCSCMTSRSNCCSVYAKCKSMCYCYGVKSMKGYVA 17168
QY 320 TCCTCAAAACCCCGGCGAGTGTGACATCGGCGGCGCGCTTTGAAGAGCTTCCCGGTG 379
Db 17169 CSYRGSMSKYCNRGSTYSTSCGCTTTTCCCCCNANTGGGGAAGCTTTTNCNKKTY 17228
QY 380 AAGAGGCCAAAGGCTTATCTCTGTTAGTGGAGCGCGCAACCAACGCGCTTATCG 439
Db 17229 YRKRGNCAMCKYNNYNNWSRSCPAGMSTCYKSSMTMSVASYCWMCMYKCSMRSA 17288
QY 440 AAGATGTGCGCGCGGCTTCTCTGTATGAAGAACCGGCCATTACCGGCCAACATCGGCTCG 499
Db 17289 SRGMSNSSYMMKMSWRMSCTMKCCWCMKVCYMMRSMRSMYSYMTASWKSRRG 17348
QY 500 TGGCTCTGAACTGCGCTTGGGTGCGGAAACGGCTTTCATCGGGGCGACGCTTCTCATC 559
Db 17349 CTRCYWC---MSKSKYKSYMMRSMRSMKSKWMSKMSGMSASRSCYKCYKSMRSC 17405
QY 560 ACAAGAAATTCGATCTGTTGAAGGCAACGATATCATTTCTGCGCTTGGCTTGGCGCTG 619
Db 17406 MMSKCYRCAGCMWKGGMVYMRWMSKMRWMSKMSMYRMRMRWKGGMWCMKCYSR 17465
QY 620 GCGATCCGACAGCTTACGCTTCGAGGTGCTCGGGGCAAGCCAAAGGTGTTAG 679
Db 17466 MSKCMWMSKSYCASRSCAMSGTNYKMSYCMWMSYTCWCTSTYMSYRCTCWCK--G 17523
QY 680 ATCTGCTTGGCGAGCATTCGCTGCCGTTTTCGAGGCAATGATCGAAGGCAACCTGCAA 739
Db 17524 WSSYTKSKSWSSSSSYKKGRRKSYSMCTCTSRGAMSCWRRCCVMRGASSMRAGSMRRA 17583
QY 740 TCATGCGCAACATTCGACAGCGGCTGCCATCAGCATATCGATCTGGCTGCCGCGG 799
Db 17584 KGRSWGGRSKWMTTGGWMSKYYYCTGRMMMTYMCWRRRSMYRSMAMGRKSS 17643
QY 800 TACTCAAGTTCGAAACGCTGCCGAGTCCCGAGCTCACGAGCTTCGAATGCTAGTCCCG 859
Db 17644 WSGRMMGASRSPRCKSASRSCWRSKMRGSCWSSKMWGSSRSRSCSKGSRGXR 17703
QY 860 TCATGAGACCGACCTCTTCTTCCAGCGCTGGGAAAGTCTGAGCTCAAAGGAAG 919
Db 17704 RRSKRSKYRGRGKRRMTKSKGSKGSKWKRSGMTSSCYYSASSCMMWSSSKSC 17763
QY 920 CGGACAGCTCTTCGAGCGAGAAATCCAGAAAAATGGGTAGGCTTCGCGCTTAACGGCT 979
Db 17764 MCOMMKRCCACYSMSCTSNYRCWGMGMSKSYSTCGTCTCKCTGKSRWYMWYSW 17823
QY 980 TCAATGATGACGACATCATGGCAC 1003
Db 17824 TSNWYRAACWYCYNSTRRMNC 17847

Search completed: June 19, 2004, 06:08:07
Job time : 4598.53 secs